TABLE OF CONTENTS

1. SPACE VEHICLE CONTROL APPLICATIONS

Design Development of the Apollo Lunar Module
Kenneth J. Cox
Manned Spacecraft Center 1 1/A10

Design Development of the Apollo Command and Service Module Thrust Vector Attitude Control System
William H. Peters
Manned Spacecraft Center 63 1/E3

Orbiting Astronomical Observatory (OAO-A2) Stabilization and Control Subsystem
Thomas E. Huber
Goddard Space Flight Center 83 1/G10

Applications of the Hybrid Coordinate Method to the TOPS Autopilot
G. E. Fleischer
Jet Propulsion Laboratory 127 2/D7

Steering Laws for Double-Gimbal Control-Moment Gyros
Stephen W. Winder
Marshall Space Flight Center 157 2/F13

2. AIRCRAFT GUIDANCE, CONTROL AND HANDLING QUALITIES

Optimum Horizon Guidance Techniques for Aircraft
Heinz Erzberger and Homer Q. Lee
Ames Research Center 175 3/A8

Inertially-Aided, All-Weather Landing Systems for C/TOL.
Robert J. Pawlak
Electronics Research Center 185 3/B4

An Approach to the Determination of Aircraft Handling Qualities Using Pilot Transfer Functions
James J. Adams and Howard G. Hatch, Jr.
Langley Research Center 199 3/C4

A Frequency Domain Approach to Handling Qualities Design
William A. Wolovich
Electronics Research Center 217 3/D8

3. SYSTEM SIMULATION AND IDENTIFICATION

Optimal Regulator or Conventional? Setup Techniques for a Model Following Simulator Control System
Dwain A. Deets
Flight Research Center 237 3/F1

TABLE OF CONTENTS (CONT'D)

3. SYSTEM SIMULATION AND IDENTIFICATION (CONT'D)

Quasi-Optimal Control of a Moving-Base Simulator
Elwood C. Stewart
Ames Research Center 253 3/G3

Some Modern Control Techniques for Human Operator Modeling and Identification
Richard S. Shirley
Electronics Research Center 261 3/G11

4. ENGINE CONTROL AND ADVANCED PROPULSION TECHNIQUES

Normal Shock and Restart Controls for a Supersonic Airbreathing Propulsion System
George H. Neiner, Gary L. Cole, and Francis J. Paulovich
Lewis Research Center 299 4/D2

Optimal Control of a Supersonic Inlet to Minimize Frequency of Inlet Upstart
Bruce Lehtinen, John R. Zeller, and Lucille C. Geyser
Lewis Research Center 323 4/F1

Control System Studies for Thermionic Reactors
Richard J. Hermsen,
California State Polytechnic College
Henrik G. Gronroos
JPL, California Institute of Technology 337 4/G1

Application of Fluidics to New Control Components
Miles O. Dustin, Vernon D. Gebben, and Robert E. Wallhagen
Lewis Research Center 365 5/B7

Minimum Energy Control of a Class of Electrically Driven Vehicles
Yilmaz E. Sahinkaya
JPL, California Institute of Technology 387 5/D6

5. ADVANCED CONTROL TECHNIQUES

A New Formulation for the Epsilon Method Applied to the Minimum Time-to-Climb Problem
Lawrence W. Taylor, Jr., and Harriet J. Smith
Flight Research Center 423 5/G1

5. ADVANCED CONTROL TECHNIQUES (CONT'D)

Determination of Optimal Control Using Imbedding of the Terminal Conditions
Raymond C. Montgomery
Langley Research Center 435 5/G13

Stability of Constant Gain Systems with Vector Feedback
George L. von Pragenau
Marshall Space Flight Center 443 6/A10

Synthesis of Finite Settling Time Discrete Systems
James A. Gatlin
Goddard Space Flight Center 473 6/C12

Design and Performance of Heart Assist or Artificial Heart Control Systems
John A. Webb, Jr., and Vernon D. Gebben
Lewis Research Center 495 6/E1

Item 830-H-10

NAS1.55: 7

MAY 12 1978

NASA Conference Publication 7

COMPLETED
ORIGINAL

NAS1.55:

Fourth NASA Inter-Center Control Systems Conference

A conference held at
Boston, Massachusetts
November 4-5, 1969

NASA

510

Item 830-H-10
NAS 1.55:7

NASA Conference Publication 7

Fourth NASA Inter-Center Control Systems Conference

A conference sponsored by
NASA Headquarters, Washington, D.C.,
hosted by the NASA Electronics
Research Center, Cambridge, Massachusetts
and held at Boston, Massachusetts
November 4-5, 1969

NASA
National Aeronautics
and Space Administration

Scientific and Technical
Information Office

1978

For sale by the National Technical Information Service
Springfield, Virginia 22161
Price – $15.25

PREFACE

To those who may think it curious that a document originating in 1969 should suddenly appear in 1978:

The proceedings of the Fourth NASA Inter-Center Control Systems Conference was produced initially in a limited quantity to fill an expected number of requests and for the use of the conference participants. The conference was held at the Electronics Research Center, which NASA closed shortly thereafter; and in the course of shutting down the activities the proceedings was inappropriately tagged with a "NASA Use Only" label.

In recent years, it has become apparent that the information is of current interest to many engaged in control system research and technology. Accordingly, to better serve the technical community, the document is being published at this time.

January 1978

Carl Janow
NASA Headquarters

FOREWORD

The Fourth NASA Intercenter Control Systems Conference was sponsored by the Electronics and Control Division, NASA Office of Advanced Research and Technology. It was hosted by the NASA Electronics Research Center, Office of Control Theory and Application, at the New England Life Hall in Boston on November 4 and 5, 1969. A major goal of the conference was to present successful applications of advanced control theory. This was indeed accomplished by the excellent panel of speakers. The design and development of the Apollo lunar module digital autopilot made it abundantly clear that modern control theory lends itself well to implementation in complex control systems. The philosophy of modern control penetrates these complexities and leads to good simulation and then to flight equipment which both controls accurately and provides system flexibility to meet changing mission requirements.

During the conference many other examples were described such as electric vehicle propulsion, control moment gyros, supersonic inlet control, all-weather landing systems, aircraft handling, system modelling and simulation. The papers included in these proceedings will give the reader very useful references which should foster even broader applications of the concepts and techniques involved.

The success of the conference is due primarily to the authors who gave fine presentations and to the many who also wrote papers for these proceedings. To them I offer my gratitude for contributing to the success of the conference. I also wish to express my appreciation to the members of the program committee without whose excellent cooperation the conference would not have been possible. The members included:

Brian Doolin	Ames Research Center
Terry Carney	Electronics Research Center
Melvin Burke	Flight Research Center
Henry Price	Goddard Space Flight Center
John Scull	Jet Propulsion Laboratory
Jarrell Elliott	Langley Research Center
Aaron Boksenbom	Lewis Research Center
Kenneth Cox	Manned Spacecraft Center
Michael Borelli	Marshall Space Flight Center
Carl Janow	NASA Headquarters

Finally, I would like to thank Hugo Schuck who conducted the panel discussion; George Zames and Terry Carney who served as session chairmen; Robert Galligan who administered all local arrangements; and Dana Pierce who handled the publication of the proceedings.

George Kovatch

George Kovatch
Conference Chairman

Section 1

SPACE VEHICLE CONTROL APPLICATIONS

DESIGN DEVELOPMENT OF THE APOLLO LUNAR MODULE*

By Kenneth J. Cox
Manned Spacecraft Center

ABSTRACT

The lunar module autopilot is a first-generation digital control system design. The two torque sources available for the control function of the descent-stage configuration consist of 16 reaction jets and a slow, gimbaled, throttlable engine. This case study includes a review of the design history, the design requirements, criteria, constraints, and general design philosophy of the control system development. Comparative flight-test results derived from design testing are presented.

INTRODUCTION

In the fall of 1964, a significant Apollo program decision was made; that is, to develop a digital autopilot (DAP) for all spacecraft primary control systems. It is noteworthy that previous manned spacecraft designs (Mercury, Gemini) involved analog control system techniques; thus, the Apollo DAP represents a first-generation design development. This paper contains a case study of the design history of the lunar module (LM) primary control system. The LM DAP, with respect to design requirements and constraints, is considered to be the most complex Apollo control system in use. Hence, significant original design concepts were necessarily required in the development process.

The general purpose and motivation of this paper are to provide some insight into the problems encountered by the control system designer. In many ways, the so-called "gap" between control system theory and practice is the result of a lack of appreciation of the severe/time and cost constraints under which the control system designer is required to produce his product.

Reading this paper, one may wonder why the total development has continued during a period of approximately 4 years; he may ask why the DAP should not be designed once and be finished. Most projects of this magnitude are iterative, because the requirements sometimes change radically, because the initial design constraints are generally poorly defined, because the inherent characteristics of the spacecraft plant are not well known, and because the basic input data for control system design are frequently not available in a timely manner. It is important to realize that major design decisions must be made, rightly or wrongly, despite the lack of fundamental input information. Because of the basic factors that characteristically make the design task difficult, the designer must adopt an implicit or explicit philosophy of providing some degree of flexibility in the control design, so as to accommodate future contingencies or unexpected developmental problems.

---

*Paper first given and printed in Workshop Preprints 1969 Case Studies in System Control, University of Colorado, August 4, 1969, sponsored by the IEEE Group on Automatic Control (copyright 1969 IEEE, Report No. 69C41-AC).

The concept of performance margin will be examined in a later section, but a point to be recognized now is that most papers on control theory emphasize obtaining optimum (or acceptable) performance for nominal situations, whereas in practice, the acceptability of the total design is most often determined by performance under extreme, off-nominal conditions. Generally, establishing explicit mathematical criteria for off-nominal performance is extremely difficult; therefore, the subjective judgment of the system designer (who uses significant simulation testing programs for design validation) is essential.

A significant problem encountered in designing the DAP was the lack of effective analysis techniques for developing and evaluating the digital control system. The major design tools used were phase-plane simulation techniques in which tradeoffs and design constants were established by "cut-and-try" methods. A more colorful manner of expressing this approach is (1) "the [design] has been chosen on the basis of theoretical investigations and empirical observations."

LIST OF ACRONYMS AND SYMBOLS

ACRONYMS

APS	ascent propulsion system
CDU	coupling data unit
CM	command module
CSM	command and service module
DAP	digital autopilot
DB	deadband
DPS	descent propulsion system
GTS	gimbal trim system
ICS	interpretive computer simulators
IMU	inertial measurement unit
LGC	LM guidance computer
LM	lunar module
PIPA	pulsed integrating pendulous accelerometer

PRM	pulse-ratio modulation
RCS	reaction control system
SM	service module
TJET	time (duration) of jet firing

SYMBOLS

A_{os} = estimated offset angular acceleration

a,b,c = mass property parameters

B = intercept constant

CDU = coupling data unit angle

$E, \dot{E}$ = attitude and rate errors

I = I(M) pitch or roll moment of inertia

$I_{yy} I_{zz}$ = principal moments of inertia

K = filter weights

L = L(M) distance from hinge pin of descent engine bell to center of mass of LM

M = estimated mass

M_U, M_V = applied torque

[M] = gimbal-to-pilot increment matrix

$\left[M_{GP}\right]$ = transformation from gimbal axes to pilot axes

$\left[M_{PC}\right]$ = transformation from pilot axes to control axes

N = noise

N_j = integer

n_t = number of sample periods

P,Q,R = pilot axes

P,U,V = control axes

T = control sample period

T_c = thrust command

t = time

u_Q, u_R = urgency functions

u = trim gimbal drive command (+1, 0, -1)

V = velocity

W = Kalman filter gains

Greek Letters

α = angular acceleration

β = lag angles

$\dot{\gamma}$ = transformation angle between U/V and U'/V' axes

ΔT_C = steering sample period

ΔT_S = control sample period

$\dot{\delta}$ = gimbal drive rate of descent engine = 0.2 deg/sec

ε = PRM attitude error

θ = gimbal angle, attitude

ω = angular velocity

Subscripts

C = cycle

cg = center of mass

d = desired angular velocity

e = attitude error

G = gimbaling

I = inner gimbal angle

i = index

J = about an axis from the firing of a single jet

M = middle gimbal angle

m = index

n = index

O = outer gimbal angle

Operators

^ = estimate

˙ = first time derivative

¨ = second time derivative

˜ = measured value

$\sum$ = summation

DESIGN CONSIDERATIONS

Vehicle Characteristics

The DAP provides stabilization and control of the LM during both coasting and powered flight in three configurations – descent (Figure 1), ascent (Figure 2), and docked with the command and service module (CSM) as shown in Figure 3. During the preliminary spacecraft-design phase, many fundamental decisions were made that define (and constrain) the control system design. For the LM, three basic propulsion force and torque systems were established – reaction control system (RCS), descent propulsion system (DPS), and ascent propulsion system (APS). Characteristics that influenced the control task included the type of actuation system, the geometrical location and number of thrusters or jets (for redundancy), and the type of thrust-variation system.

The control options available to the systems designer are divided into various flight-mode categories (Table I). The RCS

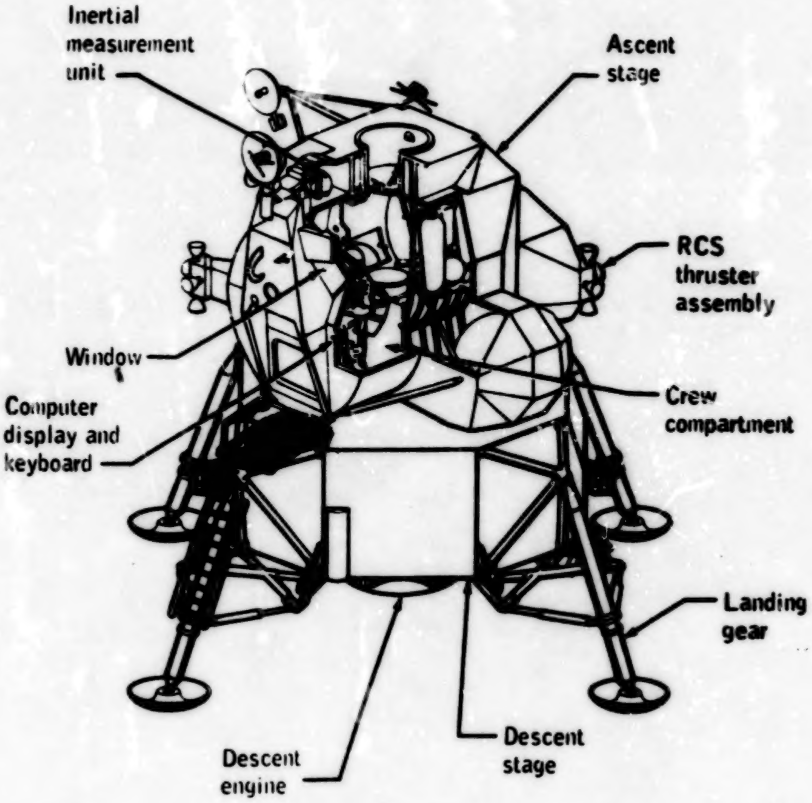

Inertial measurement unit
Ascent stage
RCS thruster assembly
Window
Computer display and keyboard
Crew compartment
Landing gear
Descent engine
Descent stage

Figure 1. The LM descent configuration

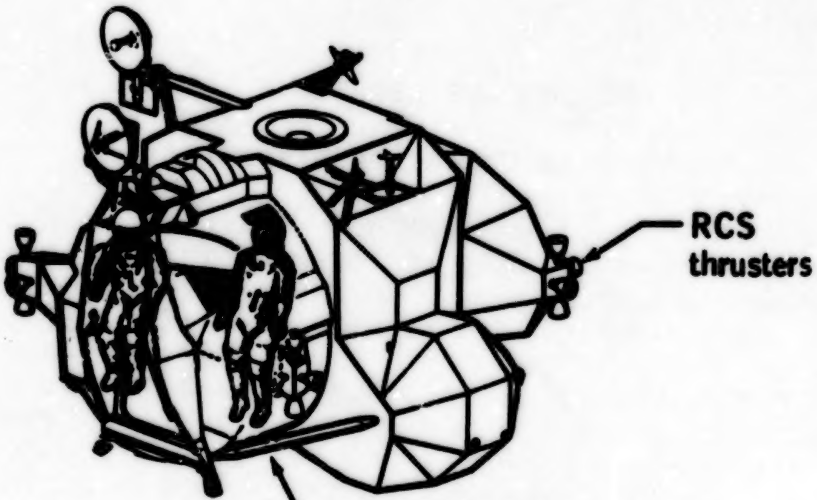

RCS thrusters
Ascent engine

Figure 2. The LM ascent configuration

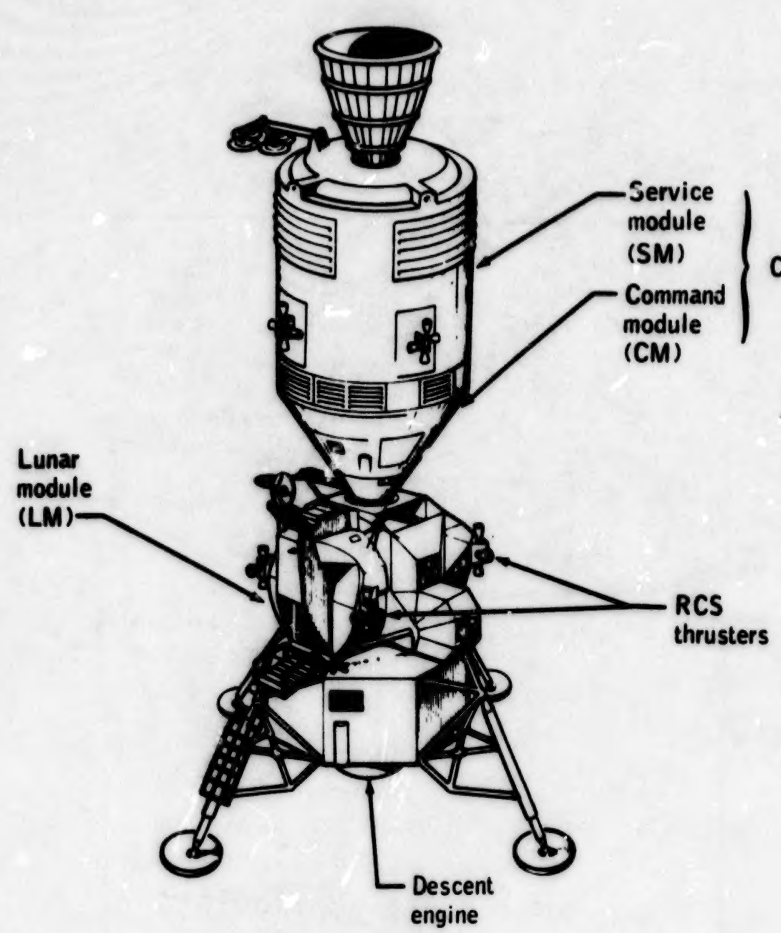


Figure 3. The CSM-docked configuration of the LM

provides automatic/manual rotation and small translation control for all LM configurations during coasting flight. During coasting flight, the design problem is characterized by the presence of extremely low disturbance torques (except for an RCS jet-on failure).

During APS-powered flight, the primary purpose of the RCS is to provide attitude stabilization and control. However, whenever feasible, it is a design requirement to fire only the upward-thrusting RCS jets to obtain ΔV in the desired direction. Because the APS is a non-gimbaled, fixed-throttle system, the RCS control laws associated with this mode must accommodate large time-variant disturbance torques.

During DPS-powered flight, the design provides yaw control with the RCS jets, and pitch/roll attitude control with a combination of the RCS and the gimbal trim system (GTS). The design

TABLE I

CHARACTERISTICS OF SPACECRAFT PROPULSION FORCES AND MOMENTS

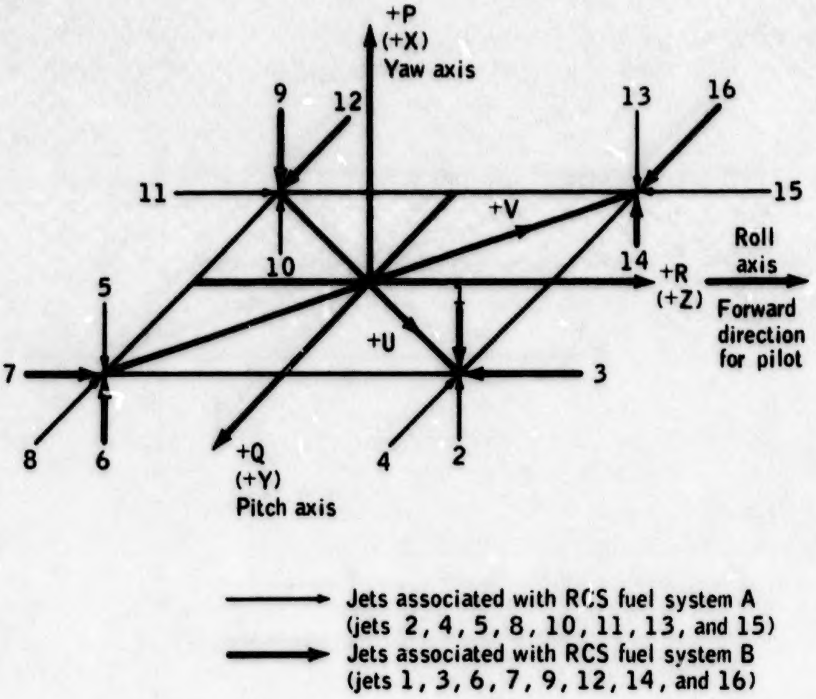
Propulsion force/moment	Vehicle configuration	Characteristics	Control function
Reaction control system	LM descent, LM ascent, LM/CSM	16 jets mounted in 4 quads 45 deg off Y/Z body axes centerline. Nominal force of 100 lb., arm length of 5.5 ft.	P-, U-, V-axis control for all configurations
Descent propulsion system	LM descent, LM/CSM	Throttable engine (1,050 to 10,500 lb.) with slow-speed gimbal actuators in Q-, R-axes	Q-, R-axis control for both configurations
Ascent propulsion system	LM ascent	Constant-thrust engine (3,500 lb.) fixed with respect to spacecraft	None

problems associated with dual control, including interactions between RCS and GTS modes, were significant for the DAP. As previously mentioned, the geometrical location of the RCS jets is significant in establishing the fundamental design approach. The locations of the RCS jets are shown in Figure 4. The eight X-axis RCS jets inherently provide control about the U/V axes, where the natural axes to consider phase-plane logic design are the Q/R pilot (or body) axes. The descent engine (not shown) is gimbaled about the pitch (Q) and roll (R) axes.

An important parameter not shown in Figure 4 is the distance from the spacecraft center of mass to the geometric center of the 16 RCS jets. This equivalent arm length is dependent upon both configuration and propellant loading, and strongly influences the ability to translate in the Y or Z direction.

Sensor Characteristics

The sensor information available for the control-design problem is provided by an inertial platform called the inertial


+P
(+X)
Yaw axis
9
12
13
16
11
15
+V
Roll
axis
10
14
+R
(+Z)
Forward
direction
for pilot
5
+U
7
3
8
6
+Q
(+Y)
Pitch axis
4
2
Jets associated with RCS fuel system A
(jets 2, 4, 5, 8, 10, 11, 13, and 15)
Jets associated with RCS fuel system B
(jets 1, 3, 6, 7, 9, 12, 14, and 16)

Notes

1. The arrows indicate thrust direction.
2. The P, Q, and R designations for the pilot axes and the P, U, and V designations for the control axes are used in connection with rotation. The X, Y, and Z designation are used in connection with translation.

Figure 4. - Locations of RCS jets on LM

measurement unit (IMU). For attitude information, gimbal angles are provided through a coupling data unit (CDU). Basic translational information is sensed by pulsed integrating pendulous accelerometers (PIPA's) located on the inertial platform. Early in the design process, the decision was made to eliminate the requirement for an independent set of rate-gyro sensors for the control function. Thus, the attitude-state-estimation function of the DAP is required to provide rate estimation, filtering (for stabilization), and disturbance-acceleration estimation (when appropriate).

Control-Mode Requirements

The DAP control modes are established primarily by mission requirements. The three required capabilities are for general attitude maneuver and attitude hold, general RCS translation, and

DPS/APS-powered-flight maneuvers. A listing of the control modes associated with the present design is presented in Table II. The design of the control modes requiring phase-plane logic will be emphasized in this case study.

Design Constraints

Numerous constraints influenced the DAP design, the most predominant class of which related to weight restrictions associated with the lunar-landing program. Weight considerations constrained the system design in structural characteristics of the LM/CSM – structural bending modes are significant; in propellant-sloshing dynamics – slosh baffles were removed early in the program; and in unbalanced couple-control requirements for APS-powered flight.

Another class of constraints, generally identified late in the design-development phase, involved restrictions on RCS jet firing. These restrictions included duty-cycle constraints (because of propulsion instabilities), exhaust-contamination constraints (particles on windows, optics), thermal constraints (rendezvous radar, antennas, spacecraft-impingement heating), and operational constraints (during extravehicular activity docking).

TABLE II

CONTROL-MODE STRUCTURE OF THE DAP

Coasting Flight	Powered Flight
Attitude hold	Attitude hold
Automatic maneuvering	Automatic steering
Manual attitude rate	Manual attitude rate
Manual X-axis rotational override	Manual X-axis rotational override
Rotational minimum impulse	Manual translation
Manual translation	

A third class of constraints that influenced the design problem was associated with propulsion-system characteristics. The slow-speed trim-gimbal-actuator characteristics of the DPS were established for crew safety to avoid hardover actuator failures during powered descent of the LM. A special gear drive was developed to restrict the trim-gimbal-drive rate to ±0.2 deg/sec. Unlike the classical actuator used for the CSM thrust-vector-control system, the DPS actuator cannot fail at a higher drive rate. A second propulsion-system constraint was associated with the decision to have a non-gimbaled APS engine. This decision imposed significant limits on allowable center-of-mass characteristics during powered ascent flight. Unfortunately, effective control of mass-property characteristics is extremely difficult in a program such as Apollo. Another propulsion-system constraint was associated with the decision to locate the RCS jets 45 degrees from the body axes. This geometry significantly influenced the interaction between the RCS mode (U/V axes) and the GTS mode (Q/R axes) during DPS-powered flight. With respect to design and development, effective analytical techniques were virtually nonexistent for this problem.

The fourth class of constraints that impacted the design problem included computer-oriented restrictions. The LM guidance computer (LGC) is limited in both fixed and erasable memories; in addition, definite timing restrictions are placed upon the programs required to provide the control functions.

Performance and Stability Criteria

The criteria for establishing the adequacy of the DAP design are outlined in Table III, which lists functional criteria for both coasting-flight and powered-flight control modes.

TABLE III

PERFORMANCE AND STABILITY CRITERIA OF THE DAP

Control modes	
Coasting flight	Powered flight
Transient behavior — acquisition and recovery	Stability characteristics
Limit-cycle characteristics	Integrated ΔV pointing accuracy
Attitude-maneuver-rate overshoot	Attitude-transient behavior
RCS propellant comsumption	Limit-cycle characteristics
Total number of jet firings	Steady-state attitude offset
	RCS propellant consumption
	Total number of jet firings

SCOPE OF PRESENTATION

Primary emphasis will be given to reviewing the design of the LM powered-flight modes. These modes are considerably more complicated than the coasting-flight modes and provide significant insight into the design principles. The manual modes, as well as the associated displays and astronaut LGC-input functions, are beyond the scope of this case study. Other areas that are not discussed include ground/spacecraft interfaces, general operational procedures, and software-implementation techniques (including coding verification).

DESIGN PHILOSOPHY

General Design Guidelines

Various decisions made in the initial DAP control-system developmental plane established general design guidelines. A partial list of these guidelines includes:

1. Simultaneous three-axis attitude maneuvering is required (as opposed to sequential-maneuver techniques).
2. Attitude rotation shall have RCS priority logic over translation (an alternate technique is a share-type logic).
3. The control system must provide acceptable performance with single RCS jet failures (on/off), with reasonable inertia-mismatch variations, and without information about single externally disabled jets.
4. The spacecraft mass-property information must be updated during DPS/APS powered-flight maneuvers.
5. The GTS control loop must be stabilized independently of the RCS control loop.

General Philosophy

Initially, an important question with respect to design philosophy was how to use the inherent flexibility associated with a spacecraft digital computer. This question was especially significant because the DAP represented a first-generation, digital design development. Emphasis was placed upon using digital capabilities, such as logic (switching, branching), non-linear computations, and function generation. For example, design considerations of the timing structure for the RCS/GTS control laws during DPS-powered flight included consideration of simultaneous control, sequential control, and time-interlaced logic control. These options are generally limited to analog control-system design. Another example was the reduction of switching-line chatter by logical branching to achieve improved performance under inertia mismatch, undetected jet failures, and ullage (X-axis translation) maneuvers.

The concept of performance margin was an area of design philosophy that influenced DAP development. This concept emphasized the principle that the acceptability of the design should be based upon performance of the system during extreme (but required) degraded conditions. Acceptable performance during off-nominal conditions, such as single undetected jet failures, and large control-effectiveness uncertainties (thrust magnitude, inertia properties, thrust misalignment, actuator drive rates, etc.) was difficult to achieve. The performance-margin concept identified a general trade-off between lowering the nominal to achieve acceptable performance during degraded conditions and maintaining high nominal performance despite severe degradation during off-nominal conditions. The control-system designer must use insight and judgment in establishing the degree of degradation (or margin) to which the design must accommodate in terms of performance. Even after this philosophy has been adopted, the ability to develop explicit mathematical criteria for off-nominal performance is still generally difficult to establish.

Another general philosophy was maximum utilization of modern control-theory techniques and frequency-domain techniques associated with sampled data-control systems. For example, the original attitude-state estimator developed for the DPS-powered flight was a Kalman filter. In addition, the GTS control loop was developed as a time-optimal control law. The many analytical methods available at that time were reviewed in reference 2, the authors of which implied that state phase-plane techniques (involving simulations) would probably be the primary design tools in DAP development. Techniques considered, but discarded, included convergence and stability (Liapunov, Aizeman, Lagrange), modified rate diagram, describing function, and dynamic programing. The concepts of defining regions of attraction and ultimately bounded regions were found to be inapplicable for this design development.

The philosophy of providing system-design flexibility to accommodate developmental problems or future contingencies was related to the concept of performance margin. An example was the guideline to stabilize the GTS control loop independently of RCS control loop. Three years after this design was initiated, additional thermal constraints (which essentially restricted all X-axis RCS jet firings) were identified for the LM/CSM configuration during DPS-powered flight. The design was flexible enough to accommodate this restriction without significantly affecting the program.

The final design philosophy listed for DAP development was associated with the RCS propellant-performance requirements. Design emphasis to achieve efficient propellant usage should be placed upon those control functions that require the largest amount of propellant over a complete mission profile. For the DAP, these control functions included manual translations, manual and automatic-attitude maneuvers, and maneuvers associated with powered-flight guidance. In this general sense, the importance of efficient RCS propellant performance for coasting-flight and powered-flight minimum-impulse operation should be downgraded. Thus, one may reasonably ask why design complexity (and associated verification) should be increased to save 20 percent performance on an item that uses 5 percent of the total mission propellant. A definite trade-off exists between design complexity and performance-improvement payoff.

DESCRIPTION OF DAP DESIGN

The DAP design that was flight-tested on the Apollo 9 manned mission will be described in this section. This design (associated with the SUNDANCE flight ropes) is described in reference 3, and will be treated as the base line design for the case study. Virtually all of the following design descriptions are condensed from reference 3.

Coasting-Flight Modes

The two coasting-flight modes described are the attitude-hold mode, and the automatic-maneuver mode. A block diagram of the coasting-flight attitude-hold mode is presented in Figure 5.

The major design elements include the attitude-state estimator, the RCS control laws, and the jet-selection logic, functional descriptions of which are provided in this section.

The basic measurements available to the recursive state estimator are the three gimbal angles from the IMU. The estimator for coasting flight predicts both attitude and angular velocity, and uses non-linear threshold logic to reject low-level measurement noise. Angular-acceleration information caused by RCS jet firings is an additional input to the state estimator.

The RCS control laws compute the requirements for rotational impulses, using information based upon attitude phase-plane errors, control effectiveness, and phase-plane targeting logic.

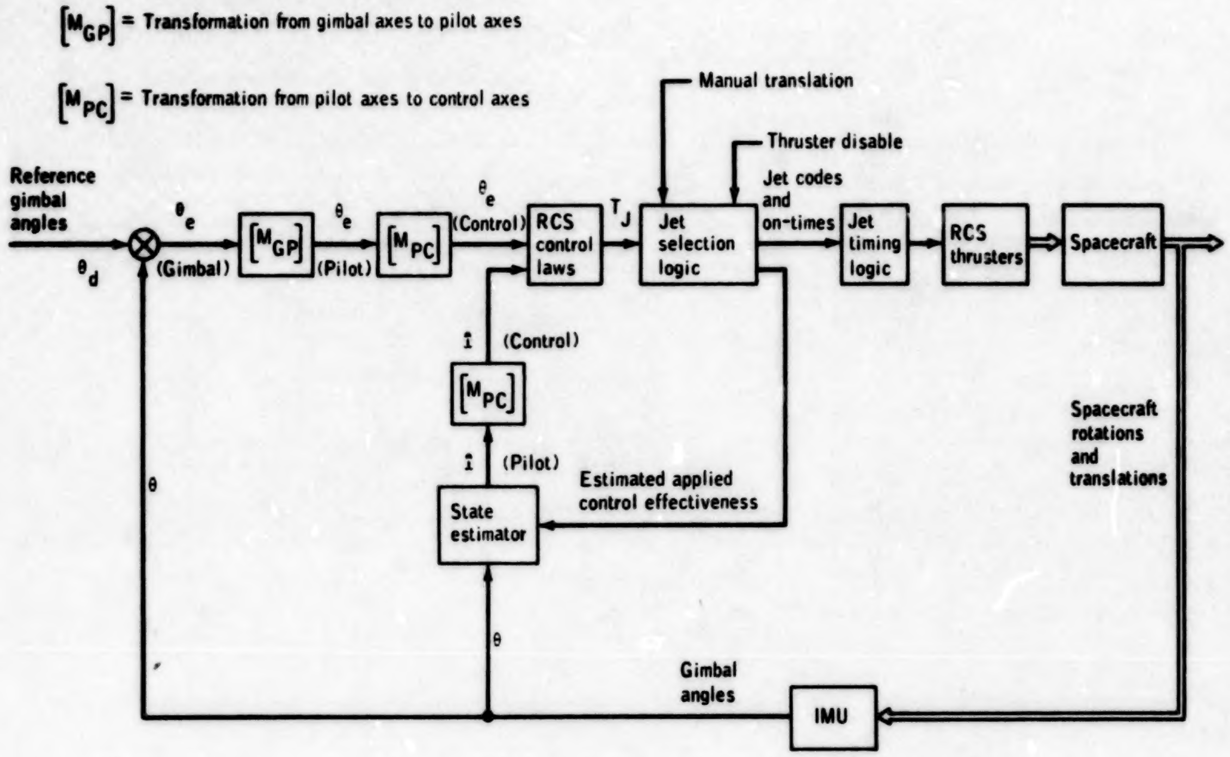

[M_GP] = Transformation from gimbal axes to pilot axes
[M_PC] = Transformation from pilot axes to control axes
Reference gimbal angles
θd
θe (Gimbal)
[M_GP]
θe (Pilot)
[M_PC]
θe (Control)
RCS control laws
T_J
Manual translation
Thruster disable
Jet selection logic
Jet codes and on-times
Jet timing logic
RCS thrusters
Spacecraft
Î (Control)
[M_PC]
Î (Pilot)
Estimated applied control effectiveness
State estimator
θ
θ
Spacecraft rotations and translations
Gimbal angles
IMU

Figure 5. - Coasting-flight attitude-hold mode

The jet-selection logic combines the required rotational impulses with the commanded translation inputs to select appropriate jets for control action. Additional information used by the jet-selection logic includes the desired number of jets to be fired and the identification of disabled jets.

Additional design elements shown in Figure 5 are transformation equations and jet-timing logic. The transformation equations, from gimbal to pilot axes, and from pilot to control axes, are straight-forward, not to be presented explicitly. The jet-timing logic is used to establish mandatory conditions for two-jet control in the U/V axes, and for four-jet control in the P-axis.

A block diagram of the coasting-flight automatic-maneuver control mode is presented in Figure 6. Automatic maneuvers are implemented using the same logic as attitude hold, except for the attitude-maneuver routine. This routine provides desired steering commands in both attitude and rate, as well as generating a set of lag angles β. These lag angles are introduced to prevent overshoot when initiating or terminating an automatic maneuver. The simplified single-plane equations for the attitude-maneuver routine are given as

$$\Delta\theta_d(t_n) = \left[\theta_d\left(t_n + 1\right) - \theta_d(t_n)\right]\frac{\Delta T_S}{\Delta T_C} \qquad (1)$$

$$\theta_d\left(N_j\right) = \theta_d\left(N_{j-1}\right) + \Delta\theta_d\left(t_n\right) \qquad (2)$$

$$\beta = \frac{\omega_d|\omega_d|}{2\alpha_J} \qquad (3)$$

where Equation (1) is computed every steering cycle ($\Delta T_C = N_j - N_{j-1} = 2$ sec), and Equation (2) is computed every control cycle ($\Delta T_S = 0.1$ sec). The value of ω_d is set by the maneuver-rate input, and α_j is defined as the magnitude of the assumed two-jet acceleration. When the maneuver is completed, ω_d, $\Delta\theta_d$, and β are reset to zero, and the control system reverts to attitude hold about the desired gimbal angles.

Powered-Flight Automatic Mode

The control operations associated with powered flight are considerably more complicated than coasting-flight operations. Major additions for both DPS- and APS-powered flight include an integrated guidance and navigation outer loop that interfaces with the DAP through a steering routine, and a mass monitor-and-control parameter routine. In addition, the state estimator is required to derive offset angular acceleration α. The RCS control laws are modified by making the control effectiveness and the phase-plane targeting logic dependent upon the estimated offset angular acceleration. During DPS-powered flight, a control law for the GTS is required. A timing-and-control-logic interaction between the RCS control and the GTS control is also required. A block diagram of the APS powered-flight automatic control is presented in Figure 7. The major design elements (state estimator, control laws, jet-selection logic) will be discussed in detail in the following sections.

Control Effectiveness

The DAP control laws and the recursive state estimator require information on the assumed RCS and GTS control effectiveness. The GTS control effectiveness is represented by the rate of change of angular acceleration, $\dot{\alpha}_G$, caused by constant angular-drive-rate command to the actuators. A flow diagram indicating those factors that relate to the GTS control effectiveness is presented in Figure 8. The factors indicated in Figure 8 are as follows:

M = estimated mass
L = L(M) distance from hinge pin of descent engine bell to center of mass of LM

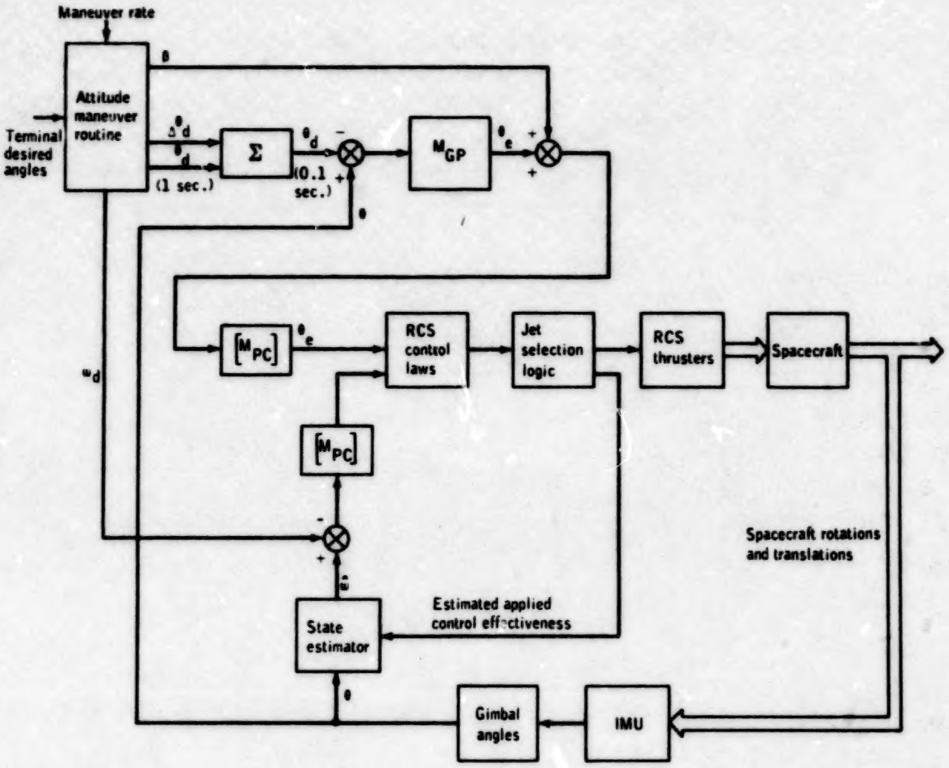

Maneuver rate
Attitude maneuver routine
Terminal desired angles
$\Delta\theta_d$
θ_d
(1 sec.)
Σ
θ_d
(0.1 sec.)
β
M_{GP}
θ_e
θ
ω_d
$[M_{PC}]$
θ_e
RCS control laws
Jet selection logic
RCS thrusters
Spacecraft
$[M_{PC}]$
Spacecraft rotations and translations
$\hat{\omega}$
Estimated applied control effectiveness
State estimator
θ
Gimbal angles
IMU

Figure 6 - Coasting-flight automatic-maneuver mode

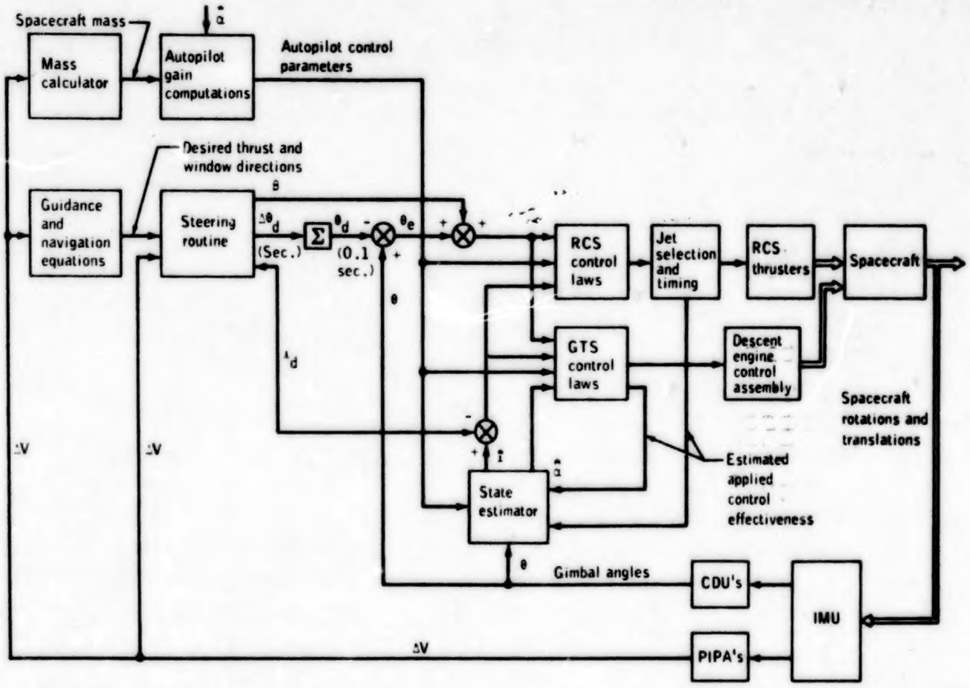

Spacecraft mass
Mass calculator
$\hat{\alpha}$
Autopilot gain computations
Autopilot control parameters
Desired thrust and window directions
Guidance and navigation equations
Steering routine
β
$\Delta\theta_d$
(Sec.)
Σ
θ_d
(0.1 sec.)
θ_e
θ
RCS control laws
Jet selection and timing
RCS thrusters
Spacecraft
GTS control laws
Descent engine control assembly
ω_d
Spacecraft rotations and translations
$\hat{\omega}$
$\hat{\alpha}$
Estimated applied control effectiveness
State estimator
ΔV
θ
Gimbal angles
CDU's
IMU
PIPA's

Figure 7. - Powered-flight automatic control of the DPS

I = $I(M)$ pitch or roll moment of inertia
$\Delta V/\Delta t$ = measured linear acceleration
u_Q, u_R = trim-gimbal drive commands (+1, 0, -1) for Q and R channels
$\dot{\delta}$ = gimbal-drive rate of descent engine = 0.2 deg /sec

The effectiveness of the RCS control is represented by the angular acceleration α_J about an axis from the firing of a single jet. A flow diagram containing the RCS control effectiveness equation in the P-, Q-, R-axes, and the appropriate equations for transformation to the P-, U-, V-axes is presented in Figure 9. To represent the RCS angular acceleration as a hyperbolic function of mass, a nominal jet-torque level of 550 foot-pounds is assumed, and additional vehicle configuration assumptions are required to establish inertia characteristics as a function of mass.

State Estimator

The recursive state estimator for powered and coasting flight is described in this section. The most complex estimator design, associated with DPS-powered flight, will be described first. Kalman filter theory provides a reasonable structure for combining estimates of state changes caused by RCS jet firings and trim-gimbal activity with external measurements of attitude. The LM plant is assumed to be represented by a simplified set of rigid body equations of motion

$$\dot{\theta} = \omega \tag{4}$$

$$\dot{\omega} = \alpha + U_J \tag{5}$$

$$\dot{\alpha} = U_G + N_{cg} \tag{6}$$

where θ = attitude
ω = angular velocity
α = offset angular acceleration caused by DPS/APS thrust
U_J = angular acceleration caused by RCS jet firings
U_G = rate of change of angular acceleration caused by descent engine gimbaling
N_{cg} = rate of change of angular acceleration caused by center-of-mass movement

The time histories of the control outputs U_J and U_G between sample instant t_{n-1} and t_n are assumed to be the values commanded, and are represented by

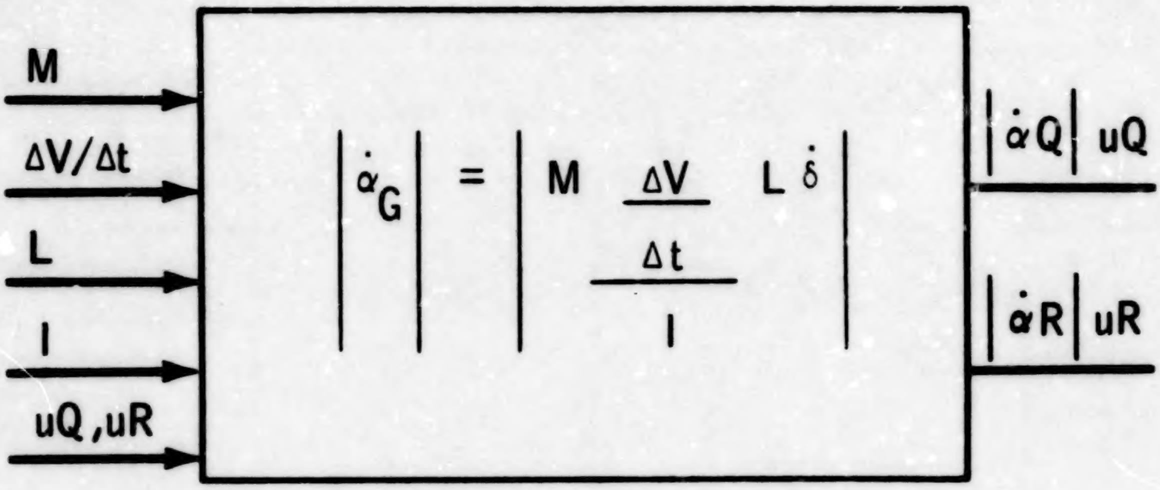

M
ΔV/Δt
L
I
uQ,uR
$\left|\dot{\alpha}_G\right| = \left|M \quad \frac{\frac{\Delta V}{\Delta t}}{I} \quad L\,\dot{\delta}\right|$
$|\dot{\alpha}Q|$ uQ
$|\dot{\alpha}R|$ uR

Figure 8. - Control effectiveness of the GTS

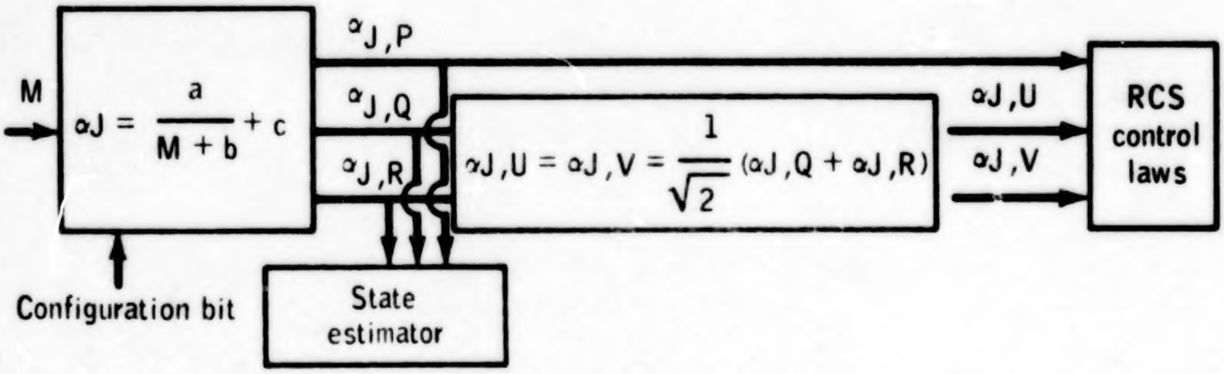
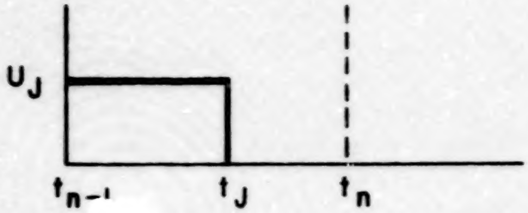
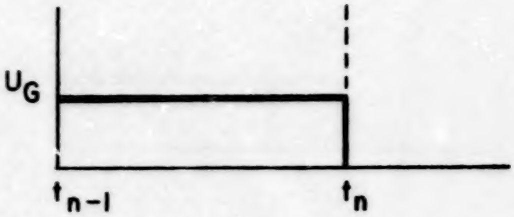

M
$\alpha J = \frac{a}{M+b} + c$
Configuration bit
$\alpha_{J,P}$
$\alpha_{J,Q}$
$\alpha_{J,R}$
$\alpha J,U = \alpha J,V = \frac{1}{\sqrt{2}}(\alpha J,Q + \alpha J,R)$
αJ,U
αJ,V
RCS control laws
State estimator

Figure 9. - Control effectiveness of the RCS





where $T = t_n$ [illegible] t_{n-1} = control-sample period.

The measurement equation is expressed by

$$\theta_m = \theta + N \tag{7}$$

where N = measurement noise from vibration, tracking errors of the CDU, and quantization. The estimation process requires the development of equations for state extrapolation and measurement incorporation. Given the estimate of the state $\hat{\theta}$, $\hat{\omega}$, $\hat{\alpha}$ at t_{n-1}, and assuming the time histories of U_J and U_G between t_{n-1} and t_n, the extrapolated state at t_n is obtained by integrating Equations (4), (5), and (6) to obtain

$$\theta'\left(t_n\right) = \hat{\theta}\left(t_{n-1}\right) + \hat{\omega}\left(t_{n-1}\right) T + \hat{\alpha}\left(t_{n-1}\right)\frac{T^2}{2} + U_G \frac{T^3}{6} + U_J t_J\left(T - \frac{t_J}{2}\right) \tag{8}$$

$$\omega'\left(t_n\right) = \hat{\omega}\left(t_{n-1}\right) + \hat{\alpha}\left(t_{n-1}\right) T + U_G \frac{T^2}{2} + U_J t_J \tag{9}$$

$$\alpha'\left(t_n\right) = \hat{\alpha}\left(t_{n-1}\right) + U_G T \tag{10}$$

The extrapolated attitude variables are compared with the external measurements to obtain an updated estimate at t_n

$$\hat{\theta}\left(t_n\right) = \theta'\left(t_n\right) + K\theta\left[\theta_m\left(t_n\right) - \theta'\left(t_n\right)\right] \tag{11}$$

$$\hat{\omega}\left(t_n\right) = \omega'\left(t_n\right) + \frac{K\omega}{T}\left[\theta_m\left(t_n\right) - \theta'\left(t_n\right)\right] \tag{12}$$

$$\hat{\alpha}\left(t_n\right) = \alpha'\left(t_n\right) + \frac{K\alpha}{T^2}\left[\theta_m\left(t_n\right) - \theta'\left(t_n\right)\right] \tag{13}$$

If the filter weights, K_θ, K_ω, and K_α, are optimized based upon Kalman filter theory, the values are time variant, and are dependent upon the values of N_{cg} and N, and the uncertainty in the initial state estimate. During the design development of the DAP, the optimum filter-gain concept was discarded after many problems were identified through simulation testing. At that time, a nonlinear threshold filter was developed as part of the base line design. This filter and the threshold values associated with the concept were established from direct engineering considerations.

Development of the nonlinear threshold filter was motivated by the specific properties of the measurement noise from the IMU. For the design considered, the predominant measurement noise is derived from the nonlinear-tracking servo characteristics of the electronic CDU's. Gimbal-angle information encoded in the LGC (for moderate angular vehicle velocities) contains high-frequency noise having a peak-to-peak amplitude of approximately 0.09 degree. It is important to note that the distribution of this noise is rectangular rather than gaussian. A trap filter using threshold logic was developed to reject this type of low-level measurement noise. The logic and associated filter-gain equations are as follows. If $|\theta_m - \theta'| \leq \theta_{max}$, then

$$K_\theta = K_\omega = K_\alpha = 0 \tag{14}$$

If $|\theta_m - \theta'| > \theta_{max}$, then

$$K_\theta = 1$$

$$K_\omega = \frac{1}{n_t + N_\omega}$$

$$K_\alpha = \frac{1}{n_t + N_\omega + N_\alpha} \cdot K_\omega \tag{15}$$

where θ_{max} = threshold value (0.14 degree)
N_ω = rate gain constant
N_α = acceleration gain constant
n_t = number of sample periods that have elapsed since the threshold was exceeded last

Extensive testing has demonstrated that the nonlinear threshold logic successfully rejects low-level measurement noise. This filter also performs well with respect to disturbances that are cyclic in nature, such as slosh and structural vibrations. The filter gains for the rate and acceleration estimates derived by

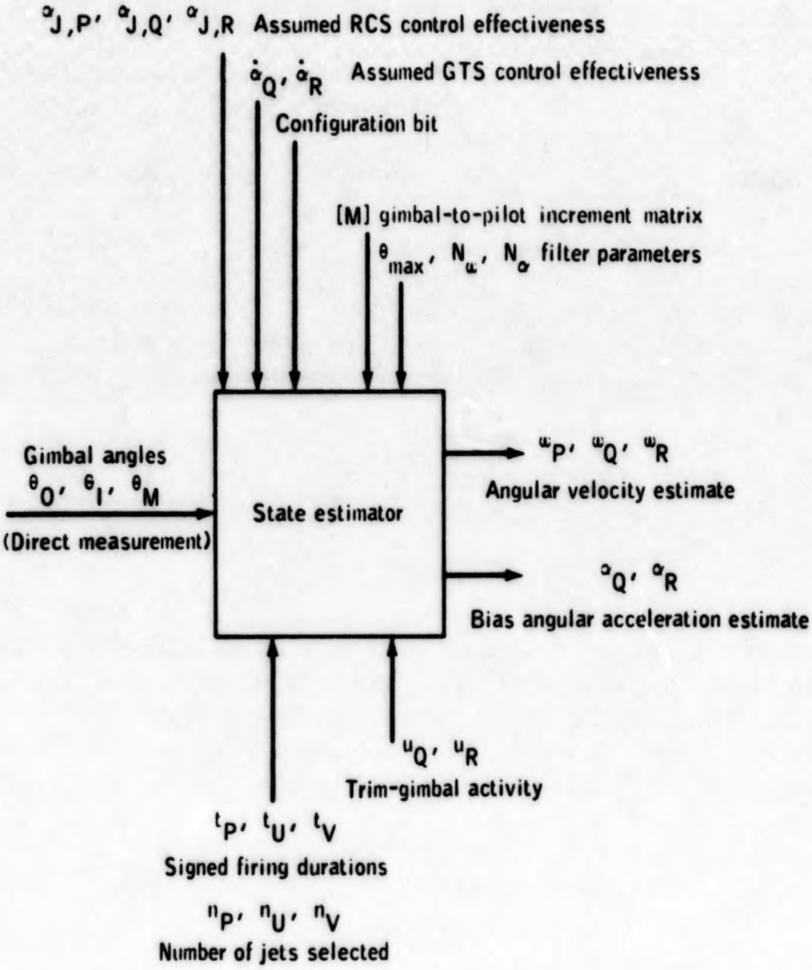
Equation (15) are functions of n_t. If the trap overflows almost every time, it is generally desirable for the filter characteristics to provide a fast rate estimate and a slow acceleration estimate. The desired response time on the acceleration estimate is set by requirements to track a moving center of mass and to respond to time-variant thrust-actuation compliance effects. The upper limit on response time is restricted because of the desire to avoid rapid fluctuations in the autopilot switching curves and because of the requirement to attenuate slosh accelerations.

If the threshold logic be exceeded only rarely, a maximum incorporation of the measurement is generally desired. The logic given in Equations (14) and (15) is actually a simplification of the developed design. The threshold value is actually compared with the total unexplained attitude that has accumulated since the last trap overflow, where the incremental amount for one control-sample period is the difference between the measured and the predicted changes in attitude.

A summary of the input and output variables associated with the state estimator for DPS-powered flight is presented in Figure 10.

The structure of the estimator for APS-powered flight is identical except that the variables associated with the GTS are deleted. Similarly, the state estimator for coasting flight is based upon the same structure, except that the estimated state does not include offset angular acceleration. The total estimator design represents an integrated concept with respect to both powered and unpowered modes of control. To conclude this section, the dynamical effects not explicitly considered in the initial development of the filter equations will be identified, as follows:

1. Propellant-slosh dynamics
2. Structural-bending dynamics
3. Jet-impingement-forces model
4. Jet-thrust lags
5. Jet-misalignment geometry
6. Jet acceleration caused by Y/Z translation
7. Undetected jet failures
8. Trim-gimbal lags
9. Inertial mismatch
10. DPS-actuator-compliance model
11. Propellant-fuel-shift model



$\alpha_{J,P}$, $\alpha_{J,Q}$, $\alpha_{J,R}$ Assumed RCS control effectiveness
$\dot{\alpha}_Q$, $\dot{\alpha}_R$ Assumed GTS control effectiveness
Configuration bit
[M] gimbal-to-pilot increment matrix
θ_{max}, N_ω, N_α filter parameters
Gimbal angles
θ_O, θ_I, θ_M
(Direct measurement)
State estimator
ω_P, ω_Q, ω_R
Angular velocity estimate
α_Q, α_R
Bias angular acceleration estimate
u_Q, u_R
Trim-gimbal activity
t_P, t_U, t_V
Signed firing durations
n_P, n_U, n_V
Number of jets selected

Figure 10. - Input and output variables of the state estimator

Other assumptions that modified the estimator equations implemented in the filter design were that the cross products of inertia terms were ignored; that the terms in Equations (8) and (9) caused by the trim-gimbal drive U_G were deleted; and that second-order rate terms were ignored in the equations of motion.

Detailed verification testing was required to demonstrate the stability and performance, including the known dynamical effects, of the total system. Simulation testing supported the establishment of the critical filter values of θ_{max}, N_ω, and N_α.

An important consideration was testing the filter performance when an undetected jet-off failure existed. When this condition occurs, the acceleration estimate will seek an average value double the actual acceleration offset present. Because the DPS acceleration-nulling control law is a function of the sign of the acceleration-offset estimate, and is invariant with the magnitude, this control law will seek the center of mass properly when undetected jet-off conditions exist.

RCS Control Laws

The control laws associated with the RCS establish jet-firing durations (TJET values) based upon phase-plane logic and assumed control effectiveness. These control laws are predictive in nature and are related to the classical two-point-boundary-value problem. To some extent, this predictive design is inherently sensitive to the uncertainties in control effectiveness and unmodeled disturbances. A description of the TJET LAWS associated with the LM descent and ascent configurations is provided in this section.

Angular error/error-rate phase planes are established for each P, U, and V RCS control axis. Because the sets of jets that produce rotation about each of these three axes are distinct, the P-, U-, and V-axis jets are turned on and off independently. A block diagram of the control-law inputs for the LM descent and ascent configurations is presented in Figure 11.

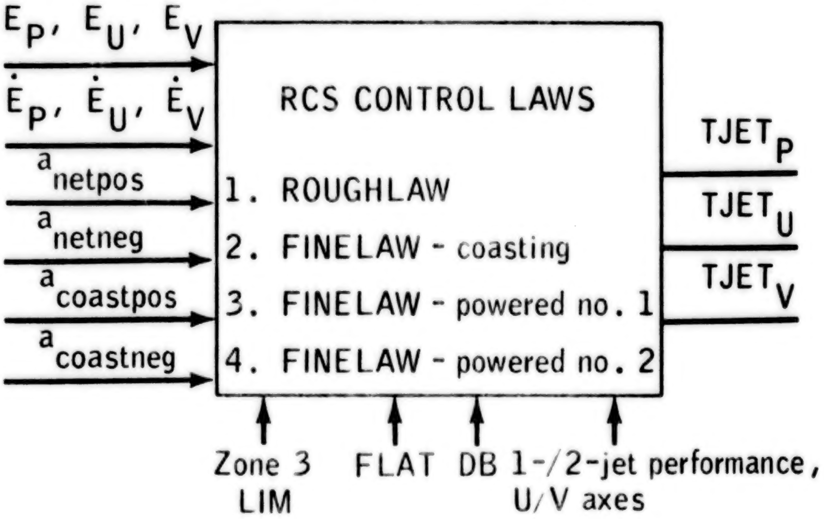
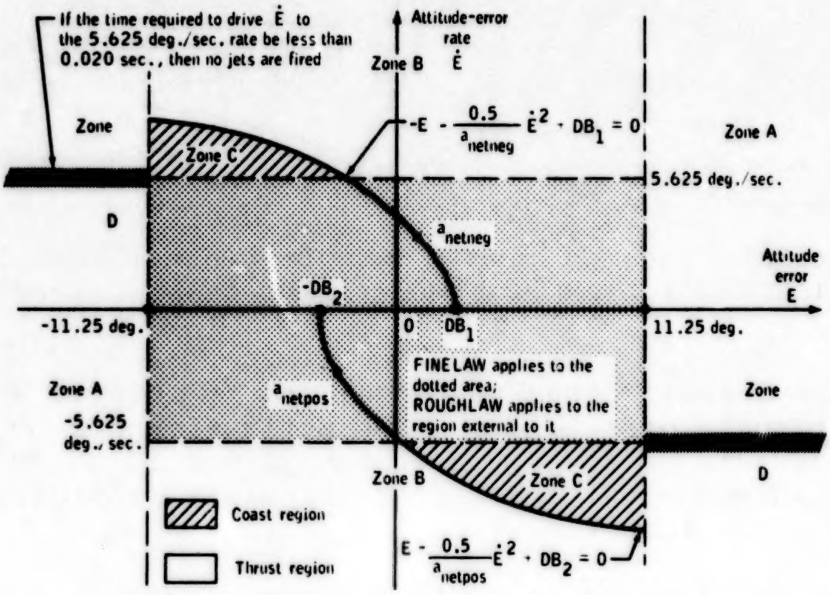

E_P, E_U, E_V
$\dot{E}_P, \dot{E}_U, \dot{E}_V$
a_{netpos}
a_{netneg}
$a_{coastpos}$
$a_{coastneg}$
RCS CONTROL LAWS
1. ROUGHLAW
2. FINELAW - coasting
3. FINELAW - powered no. 1
4. FINELAW - powered no. 2
$TJET_P$
$TJET_U$
$TJET_V$
Zone 3 LIM
FLAT
DB
1-/2-jet performance, U/V axes

Figure 11. - TJET LAWS, LM only

The attitude and rate errors $(E,\dot{E})$ are used to establish the estimated state location in the phase plane. The acceleration inputs required by the RCS control laws include net angular acceleration (jet acceleration plus offset acceleration), and coasting acceleration (acceleration caused by offset alone). The basic shape of the target parabolas and switching-line parabolas are set by this angular-acceleration information. Additional inputs – deadband (DB) and FLAT – establish the positional location of these parabolas. The four different phase-plane configurations associated with the DAP design are listed in Figure 11. The TJET control logic is developed by dividing each phase-plane configuration into coasting and firing zones.

The ROUGHLAW phase-plane configuration shown in Figure 12 is used for all control modes whenever either the magnitude of E exceeds 11.25 degrees or the magnitude of $\dot{E}$ exceeds 5.625 deg/sec. If neither of these limits be exceeded, a FINELAW phase-plane configuration is used in the RCS control logic. The use of ROUGHLAW for large values of E and $\dot{E}$ permits the efficient use of single precision arithmetic by the LGC. The computational ability to provide independent scaling for different regions of the error phase plane provides a significant flexibility compared with most of the analog control-system designs. The ROUGHLAW phase-plane configuration is divided into zones A to D. The logic for computing TJET for the upper half plane of ROUGHLAW is presented in Table IV.


Attitude-error rate $\dot{E}$
Zone B
$-E - \frac{0.5}{a_{netneg}}\dot{E}^2 + DB_1 = 0$
Zone A
Zone
Zone C
5.625 deg./sec.
D
a_{netneg}
Attitude error E
$-DB_2$
-11.25 deg.
0
DB_1
11.25 deg.
Zone A
a_{netpos}
-5.625 deg./sec.
Zone
Zone B
Zone C
D
▨ Coast region
☐ Thrust region
$E - \frac{0.5}{a_{netpos}}\dot{E}^2 + DB_2 = 0$

If the time required to drive $\dot{E}$ to the 5.625 deg./sec. rate be less than 0.020 sec., then no jets are fired

FINELAW applies to the dotted area; ROUGHLAW applies to the region external to it

Figure 12. - The ROUGHLAW phase plane

The three phase-plane configurations associated with FINELAW are shown in Figures 13, 14, and 15. The configurations include drifting flight, powered-descent flight when the trim-gimbal nulling times are less than 2 seconds, and powered-ascent flight

TABLE IV

ROUGHLAW TJET LOGIC

Location of the LM state	Basic of computing TJET
Zone A	TJET is the time required to drive $\dot{E}$ to -5.625 deg /sec.
Zone B	TJET is set equal to 0.250 sec; that is, a "large" value.
Zone C	TJET is set equal to zero, that is, no jets are turned on.
Zone D	TJET is the time required to drive $\dot{E}$ to +5.625 deg /sec.

or when the trim-gimbal nulling times are greater than 2 seconds. The FINELAW TJET logic for coasting flight is presented in Table V.

The zones 2 and 3 logic was developed to acquire a minimum-impulse limit cycle efficiently. The motivation for occasionally permitting the jets to remain on in zone 4 was to avoid switching-line chatter when the net angular acceleration is underestimated because of inertia mismatch.

The FINELAW phase-plane logic for powered-descent flight when the trim-gimbal nulling times are less than 2 seconds is identical to the coasting logic of Table V, except that zone 3 has been eliminated. This logic provides an efficient interface between the RCS and GTS control systems.

The FINELAW logic for the phase-plane configuration associated with Figure 15 is also similar to the coasting logic of Table V. The principal exceptions are that zones 2 and 3 have been eliminated, and that the logic developed from zone 4 is to target to intersect a parabola instead of the E-axis plane.

As mentioned previously, the DB_1 through DB_4 parameters, together with FLAT, are used to establish the positions of the phase-plane parabolas. The intercept values were established empirically through simulation testing, and are proportional to the DB. The proportionality constants, which were developed to depend logically upon the estimated offset acceleration, are

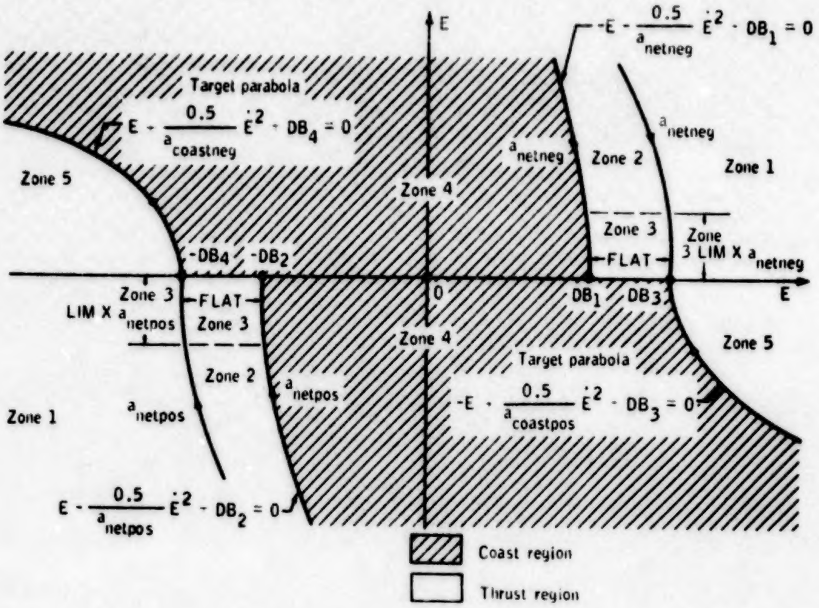

E
-E - 0.5/a_netneg Ė² - DB1 = 0
Target parabola
E + 0.5/a_coastneg Ė² + DB4 = 0
a_netneg
a_netneg
Zone 5
Zone 4
Zone 2
Zone 1
Zone 3
Zone 3 LIM X a_netneg
FLAT
-DB4
-DB2
Zone 3
FLAT
LIM X a_netpos
Zone 3
0
DB1
DB3
E
Zone 4
Zone 5
Target parabola
Zone 2
a_netpos
-E + 0.5/a_coastpos Ė² - DB3 = 0
Zone 1
a_netpos
E - 0.5/a_netpos Ė² - DB2 = 0
Coast region
Thrust region

Figure 13. - The FINELAW phase plane when the LM is in drifting flight

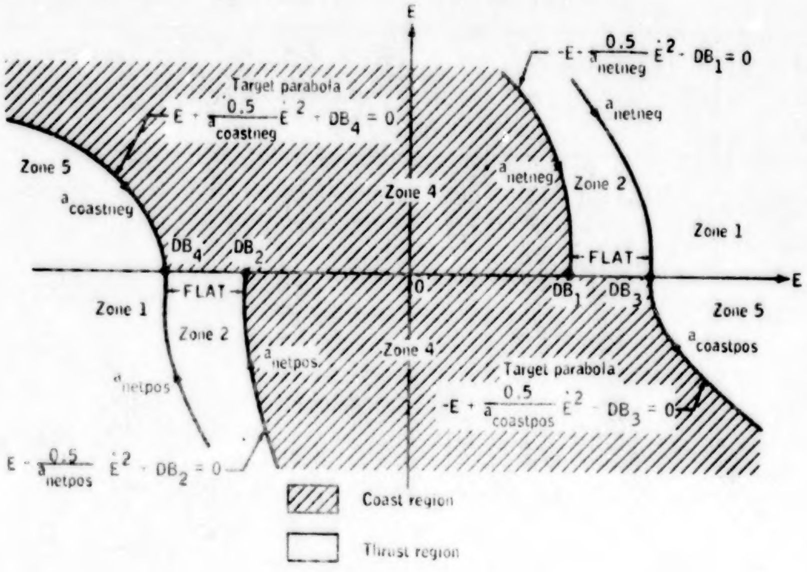

E
-E - 0.5/a_netneg Ė² - DB1 = 0
Target parabola
E + 0.5/a_coastneg Ė² + DB4 = 0
a_netneg
Zone 5
a_coastneg
a_netneg
Zone 4
Zone 2
Zone 1
DB4
DB2
FLAT
FLAT
0
DB1
DB3
E
Zone 1
Zone 2
Zone 5
a_netpos
Zone 4
a_coastpos
Target parabola
a_netpos
-E + 0.5/a_coastpos Ė² - DB3 = 0
E - 0.5/a_netpos Ė² - DB2 = 0
Coast region
Thrust region

Figure 14. - The FINELAW phase plane when both command trim-gimbal nulling drive times are less than 2 seconds during powered descent

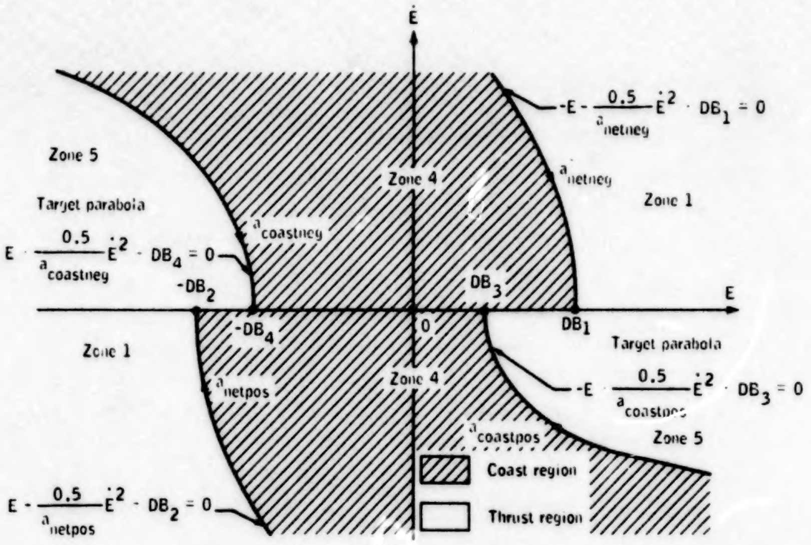

Ė
-E - 0.5/a_netneg Ė^2 · DB_1 = 0
Zone 5
Zone 4
a_netneg
Zone 1
Target parabola
a_coastneg
E · 0.5/a_coastneg Ė^2 · DB_4 = 0
-DB_2
DB_3
E
-DB_4
0
DB_1
Target parabola
Zone 1
a_netpos
Zone 4
-E · 0.5/a_coastpos Ė^2 · DB_3 = 0
a_coastpos
Zone 5
Coast region
E - 0.5/a_netpos Ė^2 · DB_2 = 0
Thrust region

Figure 15. - The FINELAW phase plane when the LM is in powered ascent or when either of the trim-gimbal nulling drive times is greater than 2 seconds during powered descent

TABLE V

FINELAW, COASTING TJET LOGIC

Location of the LM state	Basis of computing TJET
Zone 1	TJET is the time required to drive the LM state to a "target parabola."
Zone 2	TJET is the time required to drive the LM error rate to zero.
Zone 3	TJET is set so small that the jet-selection logic will fire a one-jet minimum impulse.
Zone 4	TJET is set to zero unless the jets for the axis concerned are already on and are driving the LM state to intersect the E axis between $-DB_2$ and DB_1.
Zone 5	TJET is the time required to drive the LM state to a "target parabola."

presented in Table VI. The general design strategy was to set the position of the parabolas to yield small, average, steady-state attitude errors during powered flight. The classic analog control-system trade-off of establishing a single switching-line logic to provide adequate performance during both high- and low-disturbance-torque conditions is avoided in this design.

Another input required for the RCS control laws is the indication of one-jet or two-jet couple preference for U/V axes control. Normally, one-jet control is desired for powered ascent and X-axis translation. However, additional logic associated with large phase-plane errors is used, requiring mandatory two-jet couple control when certain conditions are exceeded.

For the P-axis, the computation of TJET is always made on the basis of two-jet couples. However, the TJET logic is overridden and four jets are fired for 0.1 second if certain TJET values and error-state conditions are exceeded.

Jet-Selection Logic

The RCS control laws establish requirements for some number of jets to be fired about the P-, U-, and V-axes. The translation-acceleration requirements are obtained directly from guidance inputs or manual commands. The jet-selection logic determines the RCS jet policy when provided the desired rotational torque information, the desired direction of translation, and the desired number of jets to be used for each maneuver. In addition, the firing logic is modified when jets are known to be disabled.

The jet-selection policies associated with the P-axis rotation and the Y- and Z-axis translations are presented first. Because rotation takes priority over translation, the Y- and Z-axis commands are executed only when no P-axis commands are present. The normal P-axis jet-selection policy is presented in Table VII.

If any of the rotation policies given in Table VII involves a disabled jet, then alternate two-jet rotation policies will be attempted in the following sequential order until a policy involving only enabled jets is found.

+P Rotation	-P Rotation
7, 15	8, 16
4, 12	3, 11
4, 7	8, 11
7, 12	11, 16
12, 15	3, 16
4, 15	3, 8

TABLE VI

CRITERIA TO DETERMINE THE INTERCEPT OF THE PHASE-PLANE PARABOLAS

Define

$a_{min} = 1.4 \text{ deg /sec}^2$

A_{os} = estimated offset angular acceleration

For powered flight (except when both of the commanded trim-gimbal nulling drive times are less than 2 sec during powered descent)

FLAT = 0 and

a) If $A_{os} > a_{min}$,

$DB_1 = DB, DB_2 = 2DB, DB_3 = -0.75DB, DB_4 = 2DB$

b) If $A_{os} < -a_{min}$,

$DB_1 = 2DB, DB_2 = DB, DB_3 = 2DB, DB_4 = -0.75DB$

c) If $a_{min} > A_{os} > 0.5a_{min}$,

$DB_1 = DB, DB_2 = DB, DB_3 = 0.5DB, DB_4 = DB$

d) If $-0.5a_{min} > A_{os} > -a_{min}$,

$DB_1 = DB, DB_2 = DB, DB_3 = DB, DB_4 = 0.5DB$

e) If $0.5a_{min} > A_{os} > -0.5a_{min}$,

$DB_1 = DB_2 = DB_3 = DB_4 = DB$

For drifting flight (and when both of the commanded trim-gimbal nulling drive times are less than 2 sec during powered descent)

$DB_1 = DB_2 = DB, DB_4 = DB_3 = DB + FLAT$,

and FLAT = 0.8 deg

TABLE VII

P-AXIS NORMAL JET-SELECTION POLICY

Type of rotation	Jet-selection policy
4-jet, +P	4, 7, 12, 15
2-jet, +P	Alternate pulses between 4 and 12, and between 7 and 15
4-jet, -P	3, 8, 11, 16
2-jet, -P	Alternate pulses between 3 and 11, and between 8 and 16

The normal Y- and Z-axis translation policies, with alternate policies for disabled jets, are presented in Table VIII.

The jet-selection policies associated with the U- and V-axis rotations and the X-axis translation are now presented. If no conflict exists between jets required for translation and rotation, then both policies are executed. However, if a conflict exists, only the rotation policy is applied.

The U- and V-axis rotational policies for both normal and disabled-jet conditions are presented in Table IX.

Finally, the X-axis translational policies for both normal and disabled-jet conditions are presented in Table X.

If no alternate disabled-jet policies be possible, a computer program alarm is lighted and an alarm code informs the astronauts that a rotation or translation failure exists. This alarm procedure is applicable for all of the jet-selection policies presented.

GTS Control Laws

Two slow-speed actuators are used to gimbal the descent engine about the Q- and R-axes. The control modes developed for commanding these trim actuators are an attitude-control mode, and an acceleration-nulling mode.

The control law associated with the attitude-control mode has been developed to be a function of errors in attitude, rate,

TABLE VIII

Y- AND Z-AXIS JET-SELECTION POLICIES

Type of translation	Normal policy	Alternate disabled-jet policy
+Y	12, 16	If 16 has been disabled, set up the tacking policy of alternating between 12 and 3, and between 12 and 11. If 12 has been disabled, set up the tacking policy of alternating between 16 and 15, and between 16 and 7.
-Y	4, 8	If 8 has been disabled, set up the tacking policy of alternating between 4 and 3, and between 4 and 11. If 4 has been disabled, set up the tacking policy of alternating between 8 and 7, and between 8 and 15.
+Z	7, 11	If 11 has been disabled, set up the tacking policy of alternating between 7 and 8, and between 7 and 16. If 7 has been disabled, set up the tacking policy of alternating between 11 and 12, and between 11 and 4.
-Z	3, 15	If 15 has been disabled, set up the tacking policy of alternating between 3 and 4, and between 3 and 12. If 3 has been disabled, set up the tacking policy of alternating between 15 and 8, and between 15 and 16.

TABLE VIII

Y- AND Z-AXIS JET-SELECTION POLICIES (cont)

Type of translation	Normal policy	Alternate disabled-jet policy
+Z,+Y (+U)	7, 11, 12, 16	If either 11 or 12 has been disabled, use 7 or 16. If either 7 or 16 has been disabled, use 11 or 12.
-Z,-Y (-U)	3, 4, 8, 15	If either 8 or 15 has been disabled, use 3 or 4. If either 3 or 4 has been disabled, use 8 or 15.
+Z,-Y (+V)	4, 7, 8, 11	If either 4 or 11 has been disabled, use 7 or 8. If either 7 or 8 has been disabled, use 4 or 11.
-Z,+Y (-V)	3, 12, 15, 16	If either 15 or 16 has been disabled, use 3 or 12. If either 3 or 12 has been disabled, use 15 or 16.

TABLE IX

U- AND V-AXIS JET-SELECTION POLICIES

Type of rotation	Translational sense required	Normal policy	Alternate disabled-jet policy
2-jet,+U	-	5, 14	If 14 has been disabled, use 5 alone; if 5 has been disabled, use 14 alone.
1-jet,+U	+X	14	If 14 has been disabled, use 5.
1-jet,+U	-X	5	If 5 has been disabled, use 14.

TABLE IX

U- AND V-AXIS JET-SELECTION POLICIES (cont)

Type of rotation	Translational sense required	Normal policy	Alternate disabled-jet policy
2-jet,-U	-	6, 13	If 13 has been disabled, use 6 alone; if 6 has been disabled, use 13 alone.
1-jet,-U	+X	6	If 6 has been disabled, use 13.
1-jet,-U	-X	13	If 13 has been disabled, use 6.
2-jet,+V	-	1, 10	If 10 has been disabled, use 1 alone; if 1 has been disabled, use 10 alone.
1-jet,+V	+X	10	If 10 has been disabled, use 1.
1-jet,+V	-X	1	If 1 has been disabled, use 10.
2-jet,-V	-	2, 9	If 9 has been disabled, use 2 alone; if 2 has been disabled, use 9 alone.
1-jet,-V	+X	2	If 2 has been disabled, use 9.
1-jet,-V	-X	9	If 9 has been disabled, use 2.

TABLE X

X-AXIS JET SELECTION POLICY

Type of translation	Normal policy	Alternate disabled-jet policy
4-jet,+X	2, 6, 10, 14	If either 2 or 10 has been disabled, use 6 or 14 If either 6 or 14 has been disabled, use 2 or 10.

TABLE X

X-AXIS JET SELECTION POLICY (cont)

Type of translation	Normal policy	Alternate disabled-jet policy
2-jet,+X (fuel system A)	2, 10	If either 2 or 10 has been disabled, use 6 or 14.
2-jet,+X (fuel system B)	6, 14	If either 6 or 14 has been disabled, use 2 or 10.
4-jet,-X	1, 5, 9, 13	If either 5 or 13 has been disabled, use 1 or 9. If either 1 or 9 has been disabled, use 5 or 13.
2-jet,-X (fuel system A)	5, 13	If either 5 or 13 has been disabled, use 1 or 9.
2-jet,-X (fuel system B)	1, 9	If either 1 or 9 has been disabled, use 5 or 13.

and acceleration. The control-law equations are basically a modification of a time-optimal solution, and are given as

$$K = 0.3M \; \frac{\frac{\Delta V}{\Delta t}}{I} \; L \; \dot{\delta} \tag{16}$$

$$\Delta = -\mathrm{sgn}\left(K\dot{\theta} + \frac{\ddot{\theta}\,|\ddot{\theta}|}{2}\right) \tag{17}$$

$$u = -\mathrm{sgn}\left[K^2\theta + \ddot{\theta}\left(\frac{\ddot{\theta}^2}{3} - \Delta K\dot{\theta}\right) + \Delta\left(\Delta K\dot{\theta} - \frac{\ddot{\theta}^2}{2}\right)^{3/2}\right] \tag{18}$$

The control output commands the sign of the change in angular acceleration. The sampling period for this mode is set at 200 milliseconds. Referring to Equation (16), the time-optimal law is modified by a 0.3 gain factor in the assumed control-effectiveness term K. This reduction is designed to avoid transient-response overshoot, and to prevent large steady-state limit cycles. However, this gain should be kept reasonably high to provide good transient-response characteristics.

The GTS control law associated with the acceleration-nulling mode is designed to regulate the offset (disturbance) acceleration from the descent thrust. The primary dynamical environments that cause offset acceleration include shifting center-of-mass properties, DPS actuator compliance, and DPS engine-ablation effects. This control law is structured in the form of a trim-gimbal drive-time equation, and is given as

$$T = 0.4\left|\frac{\ddot{\theta}}{\frac{M\frac{\Delta V}{\Delta t}L\dot{\delta}}{I}}\right| \qquad (19)$$

The principal sampling period associated with the acceleration-nulling law is 2 seconds. However, under certain conditions, this acceleration-nulling law is used as part of the basic RCS control-law timing structure. The interaction and timing logic between the RCS and GTS control laws are presented in the following section.

RCS/GTS Interface

During DPS-powered flight, the GTS is generally adequate to provide satisfactory control in the Q/R axes when the maneuver requirements are slowly varying. It is believed that the GTS should provide complete control (rather than regulation of the offset acceleration) whenever possible, to limit jet firings and to minimize RCS propellant usage. During design of the DAP, a time-shared control logic was developed in which the use of RCS and GTS controls are interfaced to minimize mutual interaction.

The RCS/GTS interface has been designed so that the RCS phase-plane state is examined for a logical decision (and the RCS control law applied) at least every 200 milliseconds. The timeline operation is as follows.

t	t + 100 msec	t + 200 msec	t + 300 msec
RCS control law	Test for (1) Attitude-control law (2) Acceleration-nulling law (3) RCS control law	RCS control law	Test for (1) Attitude-control law (2) Acceleration-nulling law (3) RCS control law

The test logic associated with use of the attitude-control law every 200 milliseconds is that the trim-gimbal drive times (in the Q- and R-axes) must be less than 2 seconds [Equation (19)], and that all U and V RCS jets must be off.

The requirements associated with the use of the acceleration-nulling law are that the attitude-control law must have been used on the previous pass, and that the test logic for present use of attitude-control law must have failed.

Therefore, in the RCS/GTS timing loop, the acceleration-nulling law is used only as a transaction between the attitude-control law and the RCS control law.

In addition to the RCS/GTS timing loop, another routine executes a test for the GTS acceleration-nulling law every 2 seconds. The nulling law is applied if, and only if, the trim gimbal is not under GTS attitude control (at least one of the two test conditions is not satisfied). This "captive" logic is designed to prevent a sustained thrust offset when the RCS control laws are commanding jet firings to counteract the disturbance acceleration.

HISTORY OF DESIGN DEVELOPMENT

The history of DAP development will be discussed by presenting base line designs for the design-formulation phase (September 1964 to December 1966) and for the SUNBURST flight-program phase (December 1966 to August 1967). Where applicable, comparisons will be made to the SUNDANCE base line design previously discussed. The significant problems encountered will also be discussed.

Apollo 5, the first (unmanned) LM mission, was launched into earth orbit January 22, 1968, and used the SUNBURST flight program. Following this mission, a decision was made to simplify the DAP logic, and a significant redesign of the control system was begun, resulting in the SUNDANCE digital program. This design version was flight tested on the first manned LM mission (Apollo 9), launched March 3, 1969. Subsequent lunar-landing missions will be flown using a slightly-modified SUNDANCE flight program.

Preliminary Design Development

Many modifications in design philosophy and in control-system implementation occurred during the preliminary design phase of DAP development. Excellent insight into the various control-system problems encountered is provided by reference 4. To establish background for a discussion of design problems, a base line block diagram using assumed conditions was formulated (Figure 16). The major design areas to be discussed include RCS control-law formulation, estimation concepts, RCS/GTS control-mode interaction logic, and RCS switching logic.

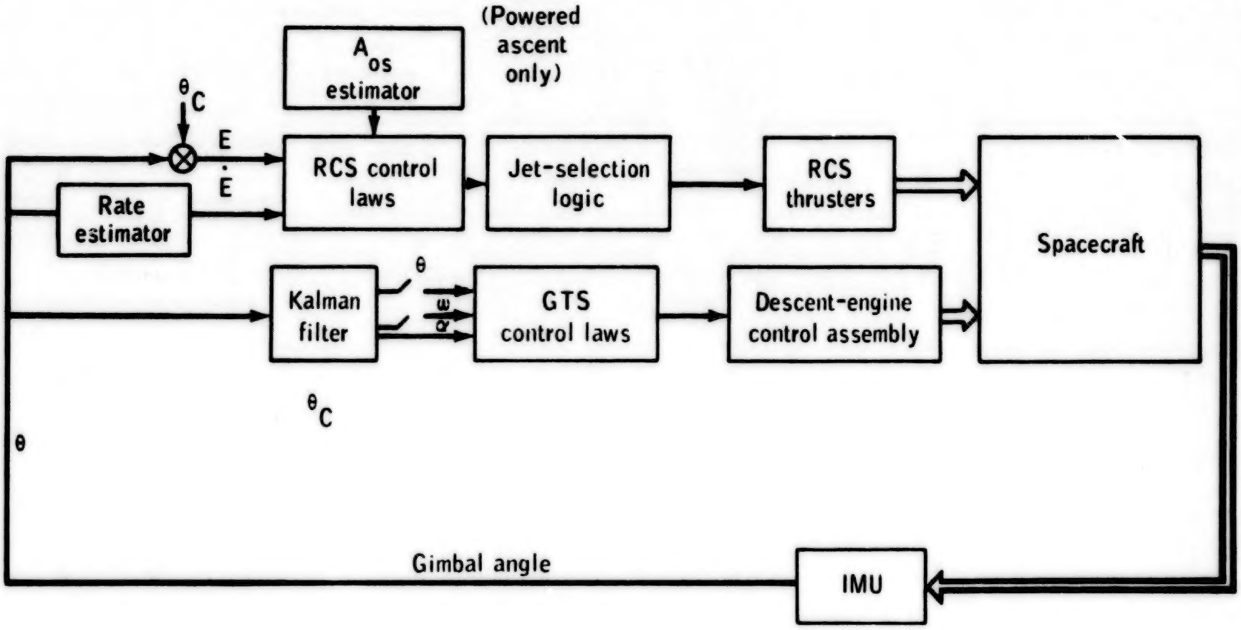

(Powered ascent only)
A_{os} estimator
θ_C
E
$\dot{E}$
RCS control laws
Jet-selection logic
RCS thrusters
Spacecraft
Rate estimator
Kalman filter
θ
ω
α
GTS control laws
Descent-engine control assembly
θ_C
θ
Gimbal angle
IMU

Figure 16. - Preliminary design control system

Three types of estimation programs were developed (an integrated design concept had not established at that time). For coasting flight, a simple rate estimator was established; for powered ascent, a combined rate and acceleration estimator was designed; and for powered descent, a Kalman filter was developed for the GTS control-law function, and a rate estimator for the RCS control-law function.

Two control modes were provided for the powered-descent operation. The primary mode controls the Q- and R-axes with the GTS, and the P-axis with the RCS. The secondary control mode uses the RCS to control all rotational axes. The primary and secondary control modes were designed to be exclusive as shown by the interface logic (Figure 17). When any primary-logic condition is exceeded, control is switched to the secondary mode, and an open-loop GTS drive is performed, using data based upon the drive time derived from Equation (19). A Kalman filter estimate of offset acceleration A_{os} is required, and is calculated in the primary mode in which undetected RCS jet failures are not a design factor. During the nominal secondary-mode operation, the RCS jets are commanded off every 10 seconds, and a Kalman filter estimate of the current offset acceleration is obtained. After this sequence, the RCS control is reinstated, accompanied by a new trim-gimbal open-loop drive. The logic conditions necessary to effect transfer from the secondary to the primary control are also presented in Figure 17. It should be noted that all conditions must be satisfied to return from secondary to primary control. The problems associated with this logic-design concept will be discussed in a following section.

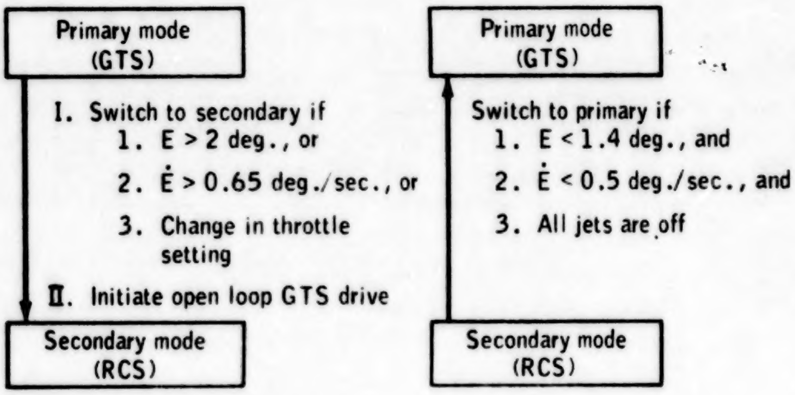

Primary mode (GTS)
I. Switch to secondary if
1. E > 2 deg., or
2. $\dot{E}$ > 0.65 deg./sec., or
3. Change in throttle setting
II. Initiate open loop GTS drive
Secondary mode (RCS)
Primary mode (GTS)
Switch to primary if
1. E < 1.4 deg., and
2. $\dot{E}$ < 0.5 deg./sec., and
3. All jets are off
Secondary mode (RCS)

Figure 17. - Interface logic of the RCS/GTS

A significant design decision required in the initial developmental period was associated with the philosophy of RCS control law. The two fundamental concepts considered were a predictive control law based upon a two-point-boundary-value approach in the error phase plane, and a logic-determination technique requiring only present- and past-state information to calculate modulated jet commands. When the two concepts were being considered, the logic-modulation technique had the advantage of considerable design experience because of analog control-system development. Hence, the basic decision was whether to establish the control-law design by digitizing a known analog-autopilot concept, or by developing a predictive control law solely based upon digital principles.

The logic-determination techniques available included pulse-frequency modulation, pulse-width modulation, pulse-ratio modulation (PRM) delta modulation, and integral pulse-frequency modulation. The development of two proposed designs that use PRM techniques is discussed in references 5 and 6. The input to the modulator is typically obtained from attitude-error and rate information, as shown in Figure 18.

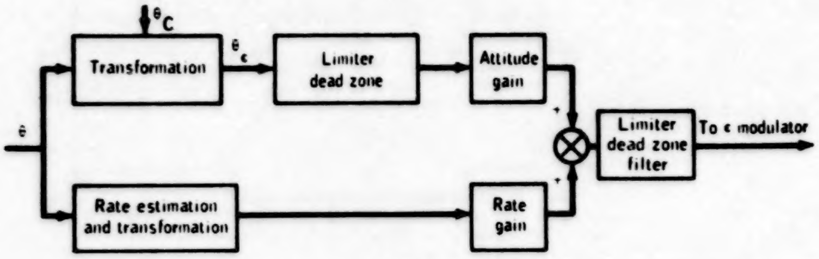


Figure 18. - Generation of control-error signal

The various trade-off factors between the digital PRM and the predictive control law are worthy of discussion. For nominal conditions, the predictive control law is generally more efficient in RCS propellant usage, (an I_{sp} penalty occurs for pulses shorter than 80 milliseconds), and usually commands a smaller number of jet firings. The principal disadvantages of the predictive approach include the sensitivity to plant uncertainties, such as inertia, thrust, and undetected jet failures, and the storage requirements for a large computer memory (parameter tracking, prediction logic, and recursive-filter techniques).

The most significant advantage of the digital PRM approach is that all logic is based upon present- and past-state information. Thus, for large off-nominal conditions, this approach has distinct advantages over the predictive design. The disadvantages include the sensitivity to noise because of low-value threshold logic, and the large steady-state attitude offsets for sustained disturbance-torque conditions. The digital PRM system cited in reference 5 estimated a sampling-rate requirement of 30 to 40 samples per second. A modification to this PRM concept, in which both on-time and off-time were calculated and the sampling requirement was reduced to 10 samples per second, is discussed in reference 6. A general trade-off exists in the area of sampling, because a good predictive system will generally require lower sampling rates than a comparative logic-determination system. However, off-nominal environmental conditions (and basic plant uncertainties) tend to increase the sampling-rate requirements of a predictive system. Thus, an estimate of expected plant uncertainties and environmental conditions is important in establishing sampling-requirement trade-offs between predictive and logic-determination control laws. After extended consideration, the decision was made to develop a predictive control law for the DAP design.

The initial concept of the phase-plane switching logic is discussed in reference 7. This design concept includes a combination of parabolic and linear curves to represent the switching and targeting lines. The phase-plane logic, assuming a positive disturbance torque, is presented in Figure 19. The value of the intercept constant B is dependent upon the deadband, jet-control acceleration, and offset-disturbance acceleration. The basic purpose of establishing switching-line equations that vary with disturbance acceleration is to lower the average steady-state attitude error.

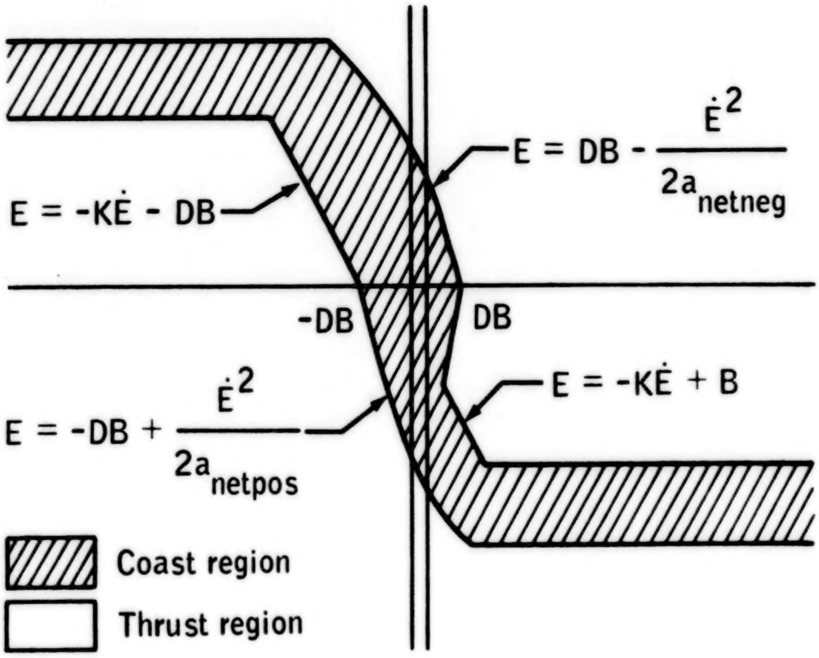

$E = DB - \frac{\dot{E}^2}{2a_{netneg}}$
$E = -K\dot{E} - DB$
-DB
DB
$E = -K\dot{E} + B$
$E = -DB + \frac{\dot{E}^2}{2a_{netpos}}$
▨ Coast region
☐ Thrust region

Figure 19. - Preliminary-design phase-plane logic

Most of the significant problems associated with the preliminary design were identified as a result of extensive simulation testing. The problem of estimating rate and acceleration when undetected jet failures existed proved to be especially difficult. Consideration was given to the use of multiple Kalman filters to estimate (from spacecraft dynamics) which of the 16 jets had failed, and to adjust the control functions accordingly. A second approach (subsequently implemented) was using the Kalman filter equations only when the GTS control law was operative or when the RCS jets were inhibited. However, disabling control during powered flight for the time needed to obtain good Kalman filter estimation was considered unacceptable and the technique was then discarded.

A second problem was that minimum-impulse operation was not achieved using the initial design. Design-verification studies indicated that this problem was caused by rate-estimation inaccuracies and quantization effects. Four phase-plane logic modifications were considered: establishing a zone 3 concept (Figure 13), discounting the computed TJET time when in zone 2, setting the derived rate (under certain conditions) equal to a relatively large magnitude with the sign changed from the value used previously, and establishing the value FLAT as a function of inertia. The first modification was implemented in the preliminary design.

Significant design problems were identified with respect to Kalman filter performance. In simulation testing, this estimator was shown to be sensitive to slosh disturbances and large initial conditions. Furthermore, during the DPS start transient, the filter performance exhibited poor convergence because of engine compliance, propellant-fuel shift, and initial engine-mistrim conditions. The manner in which the Kalman filter estimate of acceleration was initialized was also of concern. The GTS open-loop drive technique influences how the initial acceleration estimate should be set for the next pass. The filter-extrapolation equations were also modified during the preliminary design phase. Originally,, the equations did not use information on the assumed GTS control effectiveness. The addition of the U_G term [Equation (10)] substantially improved the performance of the Kalman filter.

Another preliminary design problem concerned the vehicle performance during the DPS start transient. The convergence characteristics between the primary and secondary control modes were demonstrated to be marginal. The interaction of the GTS and RCS control modes under off-nominal conditions was of concern at that time, and proved to be a major motivation in the subsequent decision to redesign the control system.

The final problem concerned rate-overshoot performance during coasting maneuvers. The command-maneuver logic did not explicitly account for the finite time required to accelerate or decelerate to the desired maneuver rate, and additional jet firings resulted. To solve this problem, lag angles were provided to prevent overshoot when initiating or terminating an automatic maneuver.

To conclude this section on preliminary design, a few general remarks on the control-system performance under off-nominal conditions are worthy of mention. Performance-verification studies indicated that the control system was insensitive to noise and small disturbance-torque conditions, but sensitive to inertia variations and thrust degradations (including undetected jet failures).

SUNBURST Design Development

The base line design developed for the initial flight program (1) and design problems that occurred in that time period are presented in this section. The flight-test results of the SUNBURST DAP, flown on the Apollo 5 mission, will be presented in a later section. This description of the SUNBURST design emphasizes the modifications and additions to the preliminary base line design.

The state-estimator equations are structured in a manner similar to that outlined in the description of preliminary design. The rate equation for coasting flight is given by

$$\hat{\omega}_n = (1 - K)\left[\hat{\omega}_{n-1} + \alpha_J t_J\right] + \frac{K}{T}\left[\overset{\gamma}{\theta}_n - \overset{\gamma}{\theta}_{n-1} + \frac{\alpha_J {t_J}^2}{2}\right] \tag{20}$$

The rate-estimation equation for powered flight is identical to Equation (20) except that the term $(1 - \frac{K}{2} T\hat{A}_{os}$ is added to the right-hand side. $\hat{A}_{os}$ is defined as the estimated disturbance acceleration caused by a main-engine thrust. During descent, $\hat{A}_{os}$ is determined every control period (T = 0.1 second) by

$$\hat{A}_{os_n} = \hat{A}_{os_{n-1}} + \frac{M\frac{\Delta V}{\Delta t} L\dot{\delta}}{I} T \tag{21}$$

For powered ascent, the A_{os} estimate is evaluated every 2 seconds by

$$\hat{A}_{os}\left(t_i\right) = C\,\hat{A}_{os}\left(t_i - 2\right) + \frac{(1 - C)}{2}\left[\hat{\omega}\left(t_i\right) - \hat{\omega}\left(t_i - 2\right) - \int_{t_{i-2}}^{t_i} \alpha_j(t)\,dt\right] \tag{22}$$

It is interesting to note that the rate- and acceleration-estimate equations are coupled for ascent. The filter gains, K and C, were established through detailed simulation testing. The nominal value of K for coast and descent is 0.5. The gains values for ascent are time variant to accommodate a nominal offset-acceleration profile and are given as

$$K = 0.44 + \frac{0.56\ t}{400} \tag{23}$$

$$C = 0.25 + \frac{0.5\ t}{400} \tag{24}$$

Kalman filter equations are used during the primary (GTS) mode and are updated every 50 milliseconds. These equations are programed in gimbal-angle coordinates and are given as

$$\mathrm{C}\hat{\mathrm{D}}\mathrm{U} = \mathrm{CDU}' + W_0\ \mathrm{C}\tilde{\mathrm{D}}\mathrm{U} - \mathrm{CDU}' \tag{25}$$

$$\mathrm{C}\hat{\dot{\mathrm{D}}}\mathrm{U} = \mathrm{C}\dot{\mathrm{D}}\mathrm{U}' + W_1\ (\mathrm{C}\tilde{\mathrm{D}}\mathrm{U} - \mathrm{CDU}') \tag{26}$$

$$\mathrm{C}\hat{\ddot{\mathrm{D}}}\mathrm{U} = \mathrm{C}\ddot{\mathrm{D}}\mathrm{U}' + W_2\ (\mathrm{C}\tilde{\mathrm{D}}\mathrm{U} - \mathrm{CDU}') \tag{27}$$

The assumed extrapolated state equations are expressed by

$$\mathrm{CDU}'_n = \mathrm{C}\hat{\mathrm{D}}\mathrm{U}_{n-1} + \mathrm{C}\hat{\dot{\mathrm{D}}}\mathrm{U}_{n-1}\ T + \mathrm{C}\hat{\ddot{\mathrm{D}}}\mathrm{U}_{n-1}\ \frac{T^2}{2} + U_G'\frac{T^3}{6} \tag{28}$$

$$\mathrm{C}\dot{\mathrm{D}}\mathrm{U}'_n = \mathrm{C}\hat{\dot{\mathrm{D}}}\mathrm{U}_{n-1} + \hat{\ddot{\mathrm{C}}}\mathrm{DU}_{n-1}\ T + U_G'\ \frac{T^2}{2} \tag{29}$$

$$\mathrm{C}\ddot{\mathrm{D}}\mathrm{U}'_n = \mathrm{C}\hat{\ddot{\mathrm{D}}}\mathrm{U}_{n-1} + U_G'\ T \tag{30}$$

where the assumed GTS control-effectiveness term is transformed to gimbal-angle coordinates.

A description of the RCS control laws associated with the SUNBURST design is necessary to the discussion of development. The basic structure of the switching lines was modified from the structure shown in Figure 19 to a format using only parabolic equations. Most of the design description has emphasized the single-plane aspects of the control-system development. A design area unique to the LM-thruster geometry (RCS jets 45 degrees from the body axes) was the logic of establishing the desired axes of rotation when simultaneous errors in pitch and roll occur. For the SUNBURST design, the Q/R axes were chosen for the control laws, and the concepts of urgency functions and urgency plane were established. Urgency functions in both the Q- and R-axes were developed to measure the state location from the coast region, as well as to measure the RCS propellant penalty if the design to apply torque be deferred. These functions are approximately equivalent to the torquing time needed to reach the boundary of the coast region. The logic flow associated with the urgency-function concept is presented in Figure 20.

The urgency plane used to select the desired axis of rotation is illustrated in Figure 21.

The two urgency functions U_Q and U_R define a position in the urgency plane and thus establish the desired axis of rotation from the eight rotation sets available (±R, ±Q, ±U, ±V). Additional logic is applied to determine the number of jets to be used in the chosen axis.

Two alternate approaches will give some insight into the design trade-offs. To illustrate the techniques, one must assume that the urgency functions in U_R and U_Q are initially established as shown in Figure 22a. The SUNBURST design logic will command a U rotation, then an R rotation, then a U rotation – until the urgency-function errors are nulled. An alternate design approach would be to null all U errors first, and then command the remaining R errors, as shown in Figure 22b.

The alternate approach is attractive in that advantage may be taken of the diagonal-jet moment arm; but, during certain disturbance-torque conditions, the delaying of the R correction could give undesirable performance. A second alternate approach (implemented in the subsequent redesign) would be to transform the original errors into the U- and V-axes, and to command U and V as shown in Figure 22c. Because the control-axis torques in U and V are decoupled, these corrections can be applied simultaneously. The logic simplifications that result from this design approach are significant; however, inefficiencies do occur for a detected jet failure, as discussed in reference 8.

A brief description of the GTS control laws associated with the SUNBURST design is needed. The time-optimal attitude-control law provided by equations (16), (17), and (18) was developed for the primary control mode in which the attitude-state errors are obtained from the Kalman filter equations. This design satisfies the requirement for an independently stabilized DPS control. The GTS acceleration-nulling law is used as part of the transfer logic from the primary mode to the secondary mode. The open-loop drive-time equation is

$$T = \frac{\hat{\alpha}}{\dfrac{M \dfrac{\Delta V}{\Delta t} L\dot{\delta}}{I}} \qquad (31)$$

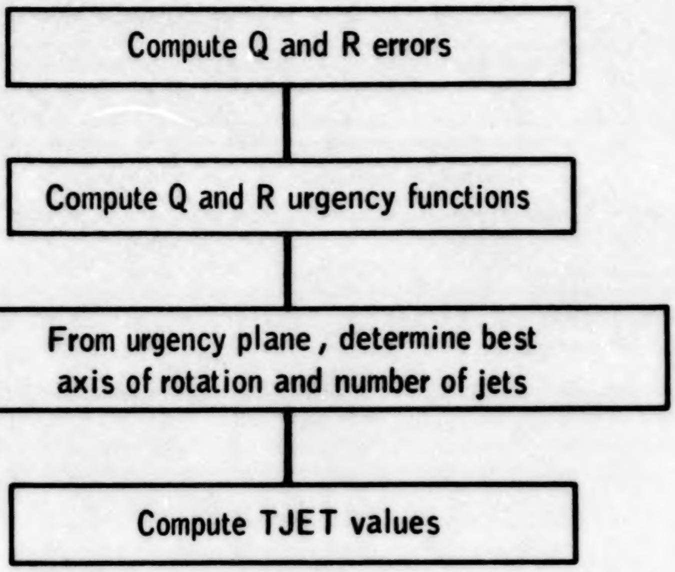

Compute Q and R errors
Compute Q and R urgency functions
From urgency plane, determine best axis of rotation and number of jets
Compute TJET values

Figure 20. - Urgency-function logic flow

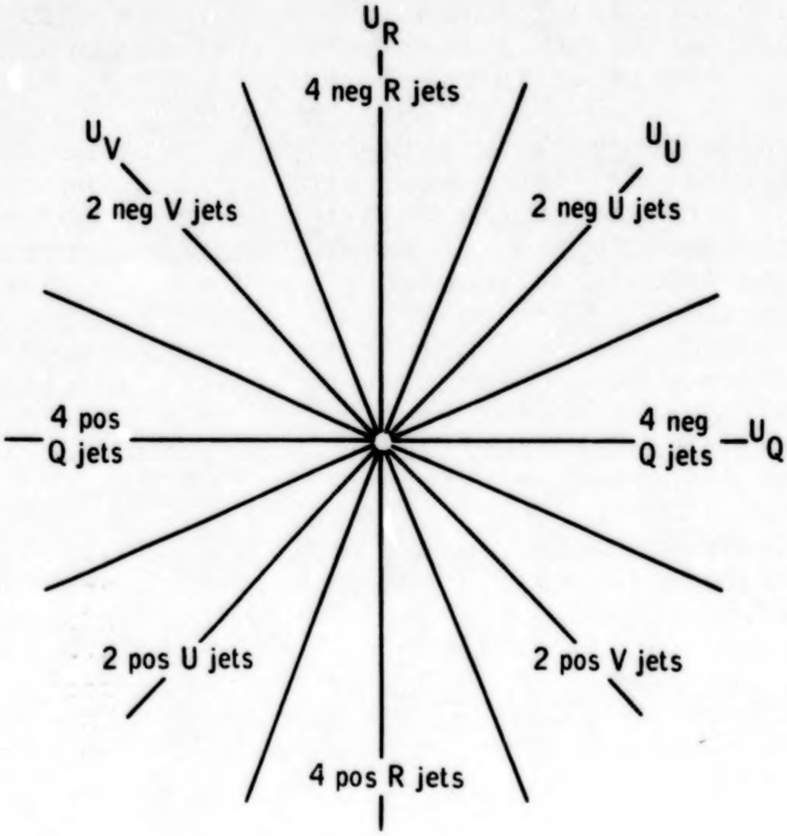

U_R
4 neg R jets
U_V
2 neg V jets
U_U
2 neg U jets
4 pos Q jets
4 neg Q jets
U_Q
2 pos U jets
2 pos V jets
4 pos R jets

Figure 21. - Urgency plane

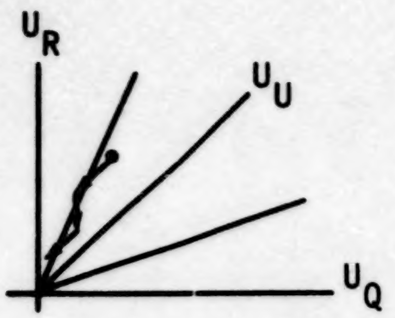


Figure 22a. - SUNBURST urgency concept

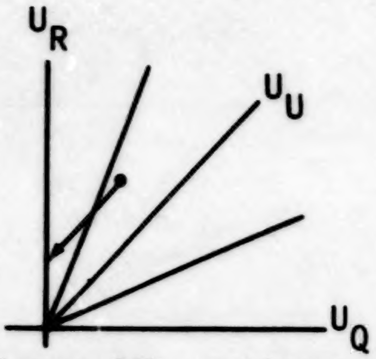


Figure 22b. - Alternate SUNBURST urgency concept

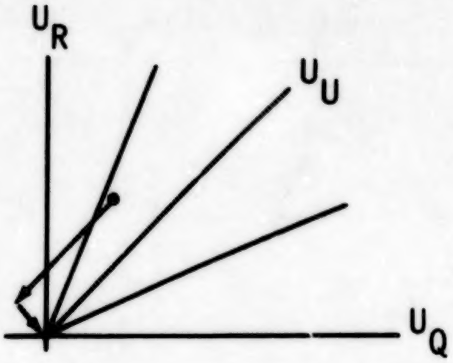


Figure 22c. - Another alternate SUNBURST urgency concept

where $\hat{\alpha}$ is the Kalman filter estimate of offset acceleration. A major difference between the SUNBURST design and the preliminary design was elimination of the technique of disabling RCS jets during powered flight to obtain Kalman filter estimation during the secondary control mode. This elimination restricted the GTS acceleration-nulling law function to an open-loop drive as part of the RCS/GTS transfer logic. The resultant performance problems associated with this interface will be discussed later.

The final SUNBURST design description to be presented are the Q- and R-axis jet-selection policies. The Q- and R-axis rotational policies for both normal and disabled-jet conditions, with alternate policies listed in order of preference, are presented in Table XI.

TABLE XI

Q- AND R-AXIS JET-SELECTION POLICIES

Type of rotation	Normal policy	Alternate disabled-jet policy
+Q	2, 5, 9, 14	2, 5; 9, 14; 2, 14; 5, 9
-Q	1, 6, 10, 13	1, 6; 10, 13; 1, 13; 6, 10
+R	1, 5, 10, 14	1, 14; 5, 10; 1, 5; 10, 14
-R	2, 6, 9, 13	6, 9; 2, 13; 2, 6; 9, 13
+U	5, 14	14; 5
-U	6, 13	6; 13
+V	1, 10	10; 1
-V	2, 9	2; 9
+Q (+X sense[a])	2, 5, 9, 14	2, 14; 2, 5; 9, 14; 5, 9
-Q (+X sense)	1, 6, 10, 13	6, 10; 1, 6; 10, 13; 1, 13
+R (+X sense)	1, 5, 10, 14	10, 14; 1, 14; 5, 10; 1, 5
-R (+X sense)	2, 6, 9, 13	2, 6; 6, 9; 2, 13; 9, 13

[a] The -X sense policies are only slight modifications of the +X sense policies listed.

Several problems occurred during the SUNBURST design phase. The general area associated with the descent primary/secondary mode interface was tested under extreme conditions, with particular emphasis upon the DPS start-transient performance. The nominal start-transient thrust profile for DPS powered-flight firings is presented in Figure 23. A "mass lockout" problem can occur for certain off-nominal conditions, when the thrust is operating at a maximum value of 10,500 pounds. One of the logic conditions needed for mandatory secondary control mode operation is that a change in throttle setting has occurred. This logic is applied when a change in thrust command T_C is sufficiently large to satisfy the inequality.

$$T_c - M\frac{\Delta V}{\Delta t} > 525 \text{ pounds}$$

When T_c is operating at a maximum value of 10,500 pounds, a mass of 5 percent or more will cause the primary control mode to be locked out. The intent of this logic was to inhibit Kalman filter estimates of offset acceleration when actuator-compliance effects were introduced by changing throttling conditions. This potential problem with the interface logic was corrected in the SUNDANCE redesign.

The performance of the primary/secondary modes during the DPS start-transient period was of sufficient concern to require design modifications before the mission. The major problem was caused by errors introduced in the open-loop drive-time equation and by the subsequent poor convergence characteristics of the primary/secondary control modes. The effect of a drive-time error is to maintain a residual offset disturbance torque while the system is in the secondary mode. If this offset be large, the RCS jets converge the attitude and rate errors very slowly to the region in which return to the GTS control is made. During this period, the jets must fire to combat the sustained offset disturbance. An example of this type trajectory behavior is shown in Figure 24.

Factors that significantly contribute to the error in open-loop drive time are

1. Propellant-fuel shift during ullage and the low throttling-time period
2. Actuator mount compliance
3. Uncertainties in the assumed values of M, L, I, and $\dot{\delta}$
4. Kalman estimate of offset acceleration
    a. Insufficient measurements
    b. Propellant-slosh dynamics
    c. Attitude-rate initial conditions
    d. Measurement noise

Simulation testing indicated that these factors could seriously degrade the performance of the control system during the DPS start-transient period. Design modifications were made to improve the RCS/GTS logic interface and the quality of the Kalman estimate of offset acceleration. The modified interface logic is given in

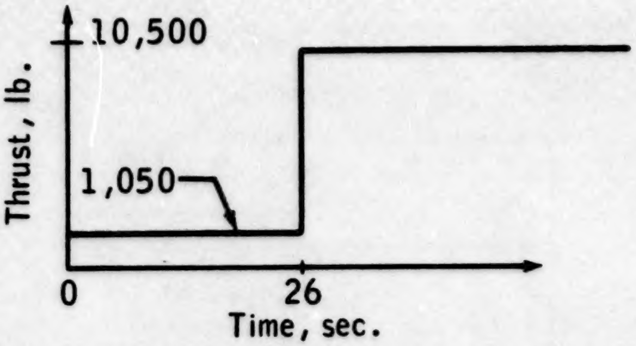


Figure 23. - Throttle profile

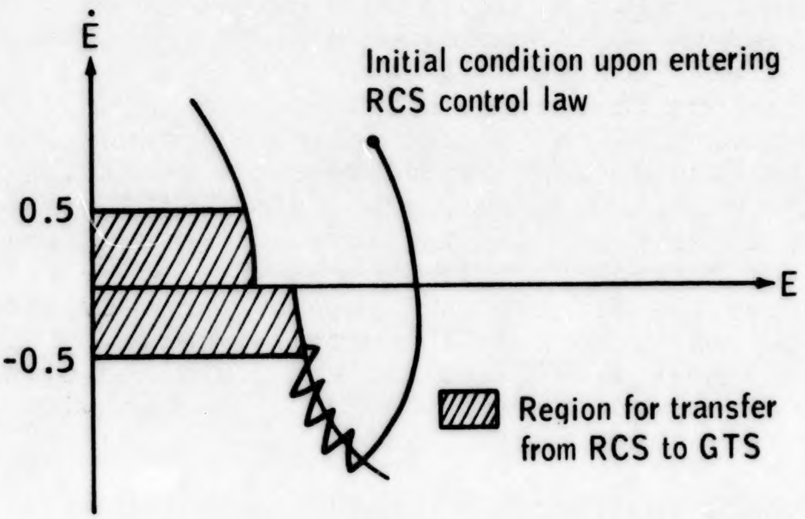

Ė
Initial condition upon entering RCS control law
0.5
E
-0.5
Region for transfer from RCS to GTS

Figure 24. - Example of RCS/GTS convergence

Figure 25. The interface logic shown provides significantly improved convergence characteristics at the expense of permitting large attitude errors during the transfer of control modes. The additional logic of insuring a minimum number of measurements for the Kalman filter was inserted because of the transient characteristic of the estimator. An actual acceleration-estimate response (9) is shown in Figure 26. For the simulation response shown, the acceleration estimate contained an error in sign for the first few measurements.

Four additional design modifications were made to improve the DPS start-transient performance: modifying the Kalman filter weighting values, limiting the maximum open-loop drive time to 15 seconds, forcing the primary control (and Kalman filter estimates) at specific times when operating in the low-throttle re-

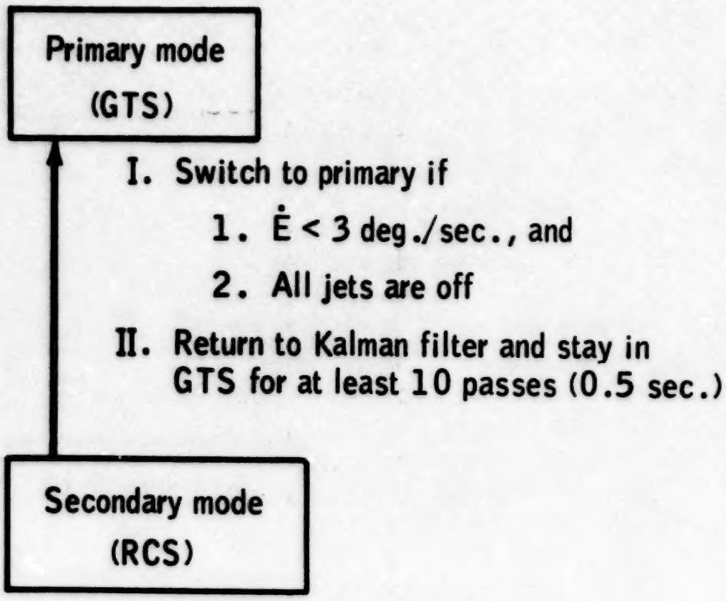

Primary mode
(GTS)
I. Switch to primary if
1. $\dot{E} < 3$ deg./sec., and
2. All jets are off
II. Return to Kalman filter and stay in GTS for at least 10 passes (0.5 sec.)
Secondary mode
(RCS)

Figure 25. - Interface logic of RCS/GTS

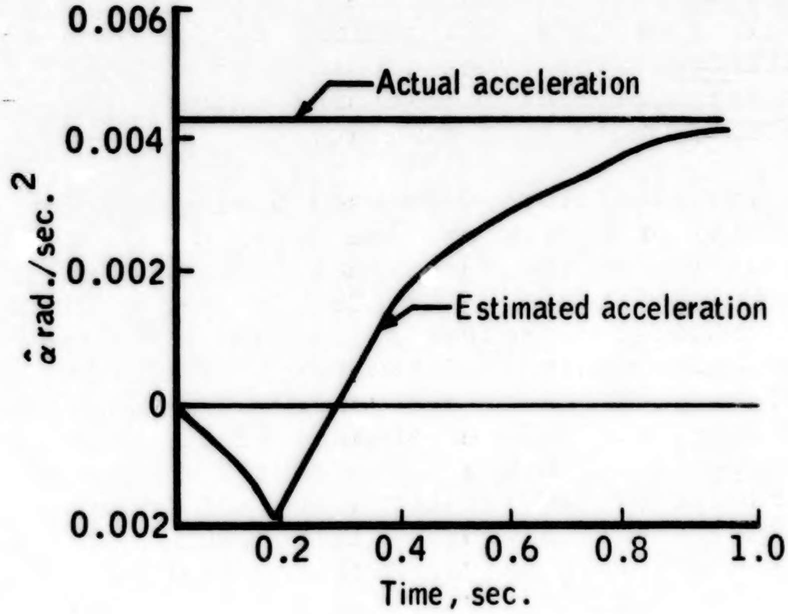


Figure 26. - Kalman filter transient performance

gion, and modifying the GTS attitude-control law during the 26-second start-transient period.

The modified GTS attitude-control law (for the start-transient period) is given by

$$u = -\mathrm{sgn}(\hat{\alpha}) \tag{32}$$

The modified design was considered acceptable for the first unmanned flight, although the inherent properties of the open-loop gimbal drive was of concern.

In August 1967, a decision was made to redesign the DAP; the SUNDANCE design, previously described, was the result of this redesign effort. The objectives of the redesign were to reduce memory-storage requirements, improve off-nominal performance, and reduce computer-execution time. The five major changes that resulted included elimination of the urgency-function concept, simplification of the jet-selection logic, simplification of the RCS control-law logic, improvement in the GTS/RCS interface design, and development of an integrated state-estimator design.

TESTING AND VERIFICATION

The mission-verification and design testing conducted on the SUNDANCE DAP is discussed next. The primary objectives of preflight testing were to validate the control-system performance during nominal conditions, off-nominal conditions, and mission-related conditions. The types of simulation facilities used included engineering digital simulators, interpretive computer simulators (ICS), and hybrid simulators.

Engineering simulators were used during initial development (or modification) phases to provide dynamic validation and performance evaluation of the functional design under a broad spectrum of mission conditions. The ICS bit-by-bit simulator modeled the detailed computer characteristics, and was used to verify the software-programing design. Parameter-type studies associated with off-nominal performance are generally inefficient to run on the ICS. However, nominal-performance verification is conducted on a mission-by-mission basis. The hybrid simulators were used to verify hardware/software interfaces, and to provide overall system-performance validation. With respect to the DAP, both design-validation and mission-verification testing programs were conducted on hybrid simulators.

The formal testing conducted on the SUNDANCE DAP design is reported in references 8, 10, and 11. Reference 10 is an excellent test-results document. All control modes of flight were tested during nominal-performance conditions, RCS jet-failure conditions, and incorrect-mass-property conditions. Powered-flight testing included recovery from large rate and attitude errors, DPS/APS start-transient performance, and performance with large offset accelerations. A general summary of the test results follows.

1. Nominal performance was satisfactory (all modes).
2. Minimum-impulse limit cycles were achieved (coasting modes).
3. Efficient, automatic attitude maneuvers were achieved.
4. Translation-acceleration capability was degraded by jet failures.
5. Powered flight modes were relatively insensitive to inertia mismatch (errors of ±25 percent were tolerable).

Several test results from reference 10 are presented to indicate performance trends.

The RCS propellant consumed during a 2-degree-per-second maneuver is shown in Figure 27 as a function of mass mismatch. The theoretical fuel (1.21 pounds) is substantially below the minimum fuel usage (1.55 pounds), because the theoretical model does not account for jet-plume-impingement forces. A summary of descent-configuration maneuver performance for various jet-failed conditions is presented in Table XII.

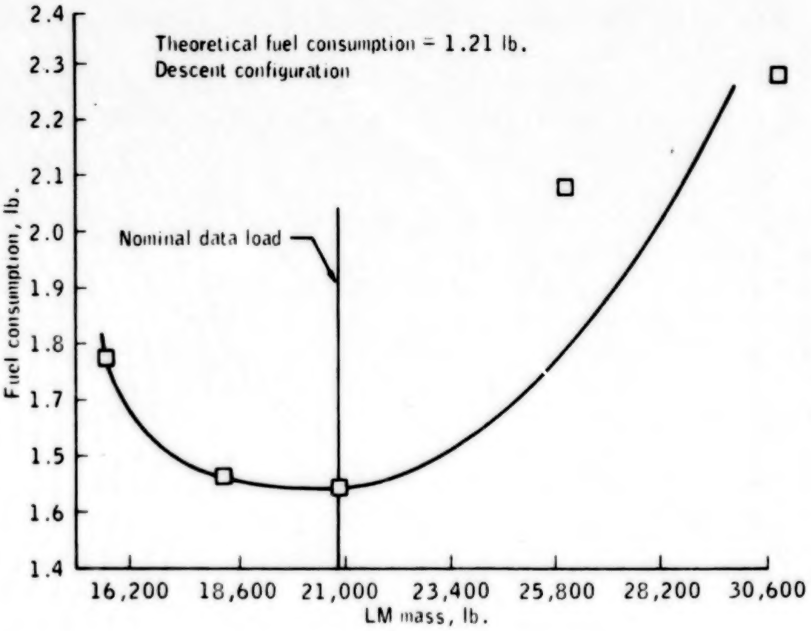


Figure 27. - RCS fuel for 2 deg /sec maneuver

For the third condition listed, the fuel consumed was less than nominal. The reason for this paradox is that the jet 10 impingement force is larger than the impingement forces associated with the other jets, so that overall system efficiency is higher if jet 10 control be deleted.

The final performance curve shown is presented in Figure 28, in which the RCS propellant-versus-maneuver rate is presented for a LM descent configuration. The relative maneuver efficiency of the DAP design is difficult to assess because the theoretical fuel consumption used as a standard does not include the effects of jet plume impingement or the effects of crossproducts of inertia. The theoretical fuel consumption includes, however, the fuel required for acceleration and deceleration, the effect of crosscoupling torques, and the minimum-impulse fuel required to maintain the angular deadbands during the maneuver.

A brief summary of the hybrid-simulation test results reported in reference 8 will be given. The control system was sub-

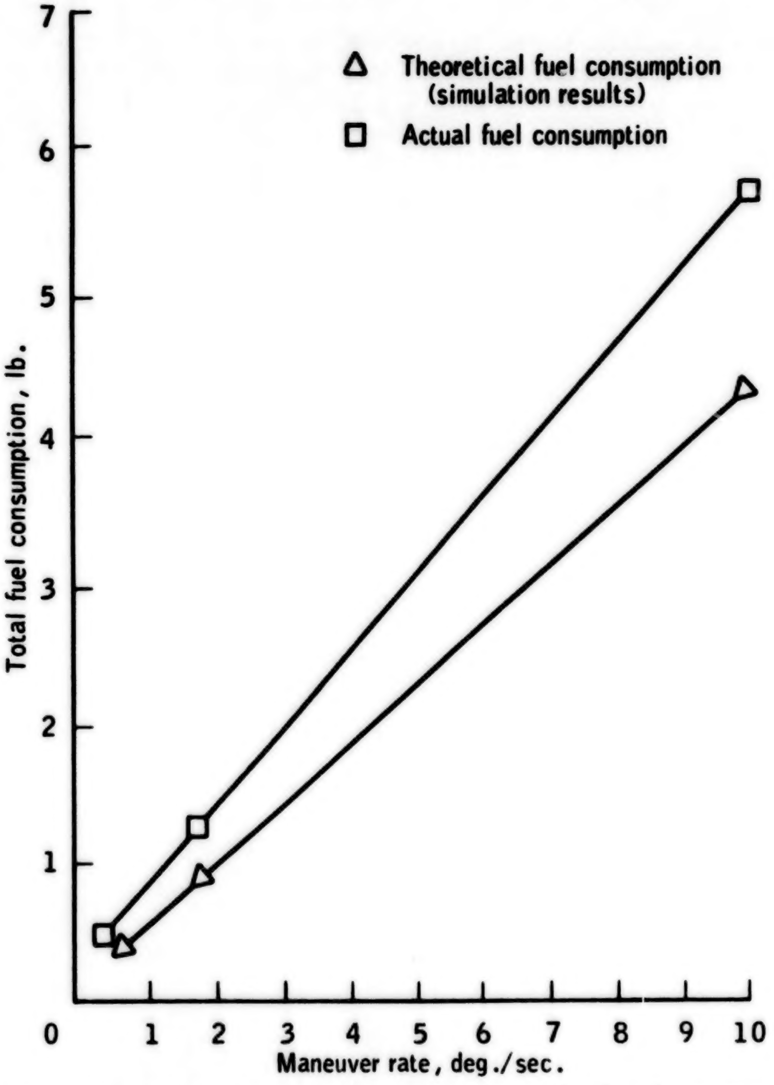


Figure 28.- Theoretical and actual RCS fuel comsumption for LM descent configuration

TABLE XII

DESCENT-CONFIGURATION MANEUVER TEST SUMMARY

Jets Failed	Maneuver rate, deg /sec		Maneuver		Fuel consumed during maneuver lb	Theoretical fuel needed to perform maneuver lb
	Desired	Greatest obtained	Desired gimbal angles	Gimbal angles at end of run		
None	2	2.04	20 80 40	19.01 81.09 39.25	1.55	1.21
Jets 10 and 11 failed off undetected	2	2.07	20 80 40	20.03 80.30 39.86	1.67	1.21
Jets 10 and 11 failed off detected	2	2.06	20 80 40	19.84 80.00 39.98	1.34	1.21
A system failed off undetected	2	2.09	20 80 40	19.48 80.86 39.47	1.83	1.21
A system failed off detected	2	2.03	20 80 40	20.17 80.00 40.48	1.55	1.21

jected to a realistic flight environment including the effects of RCS thruster impingement, propellant slosh, and actuator compliance. Off-nominal spacecraft environments included inertia, thrust, and center-of-mass uncertainties; DPS actuator-drive uncertainty; and RCS jet failures. The verification-run matrix associated with the hybrid testing was quite extensive. Integrated guidance and control hardware and a flight attitude table were used in the test facilities. Twenty-seven discrepancy items reported by the testing activity required formal disposition. Virtually all items requiring minor design modifications were incorporated into the lunar-landing-mission program.

An interesting design problem occurred in the area of inertia cross-coupling effects. With the TJET calculations established in the U/V-axes system, an RCS torque applied around the U-axis produces not only an acceleration around the desired U-axis, but also, in general, a coupled acceleration about the V-axis. The same situation applies to an RCS torque applied around the V-axis. The simplified equations of motion that demonstrate the effect of inertia crosscoupling are written as

$$\dot{\omega}_U = \frac{I_{yy} + I_{zz}}{2I_{yy}\,I_{zz}}\,M_U + \frac{I_{yy} - I_{zz}}{2I_{yy}\,I_{zz}}\,M_V \qquad (33)$$

$$\dot{\omega}_V = \frac{I_{yy} - I_{zz}}{2I_{yy}\,I_{zz}}\,M_U + \frac{I_{yy} + I_{zz}}{2I_{yy}\,I_{zz}}\,M_V \qquad (34)$$

where ω = angular velocity

M_U, M_V = applied torque

I_{yy}, I_{zz} = principal moments of inertia

This inertia crosscoupling effect between applied U/V torques and resulting U/V angular accelerations is significant only when the pitch and roll moments of inertia are substantially different. For powered-ascent operation, these inertia values were sufficient to cause crosscoupling that resulted in undesired limit-cycle performance. A subsequent design modification was made to eliminate the inertia crosscoupling effects. A nonorthogonal set of control axes U'/V' was defined which has the property that a pure U torque produces no observable acceleration in the V' direction, and a pure V torque produces no observable acceleration in the U' direction. This U'/V' nonorthogonal system can be constructed as shown in Figure 29.

The desired relationship between the U- and V-axes, and the U- and V-axes can be obtained if the transformation angle satisfies

$$\gamma = \cos^{-1} \frac{I_{yy}}{[I_{yy}^{2} + I_{zz}]^{1/2}} - 45^{\circ} \tag{35}$$

The implementation of this control law where the crosscoupling accelerations are eliminated is described as follows. The vehicle state is estimated in the P-, Q-, R-axis system. When RCS control is to be applied, the Q and R components of error angle and rate are transformed to the U'/V' system. Errors about the U'-axis are controlled by firing a U-axis RCS torque, and errors about the V'-axis are controlled by firing a V-axis RCS torque.

FLIGHT TEST RESULTS

The flight results discussed include test data from the unmanned Apollo 5 mission and the manned Apollo 9 mission. Typical flight-data results are presented to indicate performance trends. The ability to match the preflight-simulation test results closely with the actual flight data is dependent upon the quality of the telemetered data and the knowledge of the spacecraft environment. In general, powered-flight maneuvers and coasting-flight attitude maneuvers can be closely duplicated, but attitude-hold limit-cycle behavior is more difficult to match in the postflight analysis process. For the test data shown, a data-sampling frequency of one sample per second was available.

Only the DAP coasting-flight modes were exercised on the Apollo 5 flight. Flight data for an automatic 5-degree-per-second attitude maneuver showed close agreement with simulation data. As reported in reference 12, the Apollo 5 mission provided some unplanned limit-cycle data during coasting ascent because of a mass-mismatch condition. This situation arose because, although the spacecraft was actually in an ascent configuration, the DAP computed the vehicle inertia to be that of the unstaged LM. As a consequence of the 300-percent inertia-mismatch condition, a virtually continuous-firing limit cycle resulted. The narrow-deadband attitude-hold logic did maintain the desired conditions, however. After this operation, one RCS propellant system was allowed to fire to depletion, and data were taken at various lower thrust levels as the propellant was being depleted. Almost immediately, the limit-cycle characteristics began to improve, and eventually the attitude-hold function settled into a minimum-impulse condition.

Limit-cycle data were also analyzed during the descent-coast phase of the Apollo 5 mission (13). Unexplained limit-cycle trajectories in both pitch and roll phase planes, which were asymmetrical in computed error rate and symmetrical in attitude error, were observed. A representative trajectory is shown in Figure 30. During a 2-hour period, 125 jet firings occurred, approximately 30 of which had durations of from 50 to 110 milliseconds.

Preflight verification testing indicated that 16-millisecond (minimum-impulse) firings should occur at the deadband extremities. An extended effort was made to match the flight-test data through simulation testing. Inertia coupling, aerodynamic torques, and diagonal firing logic were all examined, but the observed limit-cycle phenomenon was only partially explained.

During another descent-coast phase, a different limit-cycle characteristic (Figure 31) was obtained. These trajectories contained 20-millisecond firing times, with the limit cycle restrained to one side of the attitude deadband.

This trajectory condition generally occurs during sustained torque disturbances. Limited post-flight data prevented identifying the exact nature of this disturbance, but a combination of aerodynamic torque and rate-estimation error was believed to have been the cause.

The Apollo 9 mission, during which the LM was manned for the first time, was flown in earth orbit. All powered- and coasting-flight DAP modes were exercised during the mission, and the control system performance was generally excellent. No anomalous or unexpected control-system conditions occurred. Data examined in the postflight analysis included peak-to-peak rates, attitude-deadband excursions, general limit-cycle characteristics (including existence of disturbance torques), and trim-gimbal performance.

Several flight-data results are given to indicate general performance. A 2-degree-per-second maneuver response for the ascent configuration is illustrated in Figure 32. A slight overshoot occurred in the Q- and R-axes, but overall rate performance was satisfactory. This overshoot was caused by rate-estimator errors.

A phase-plane plot of the limit-cycle performance during a powered ascent firing is presented in Figure 33. The intent of the plot is to trace the shape of the limit-cycle trajectory. Because of the data-sampling limitations, only discrete data points in the phase plane are available. The plot does indicate on a quantitative basis that the results are within a range consistent with preflight simulation results.

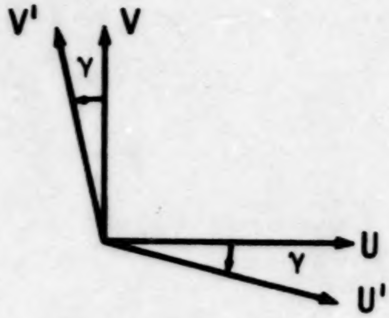


Figure 29. - Nonorthogonal U'/V' system

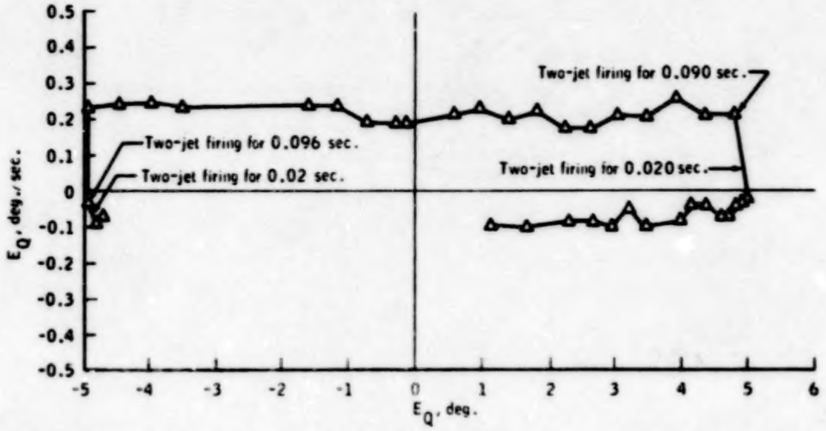


Figure 30. - Pitch-rate error versus pitch error during a descent-coast phase

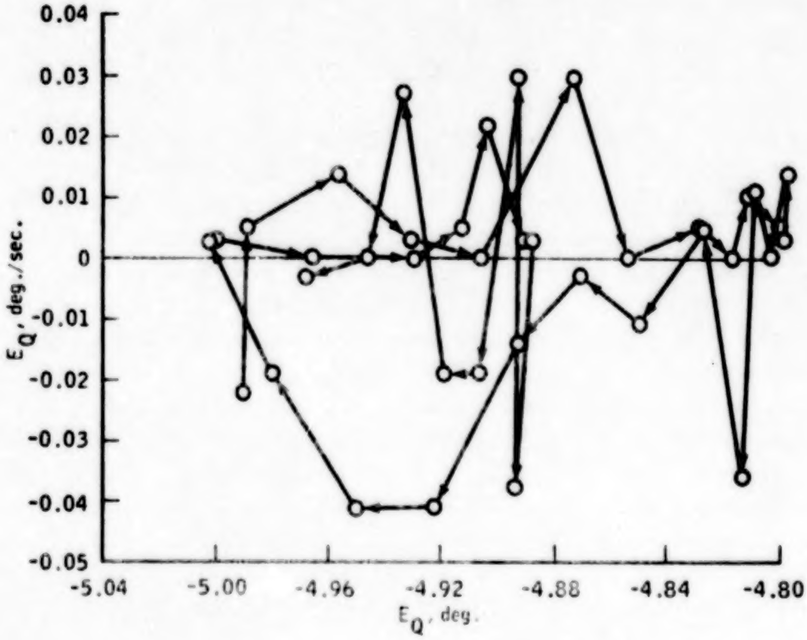


Figure 31.- Pitch-rate error versus pitch error during another descent-coast phase

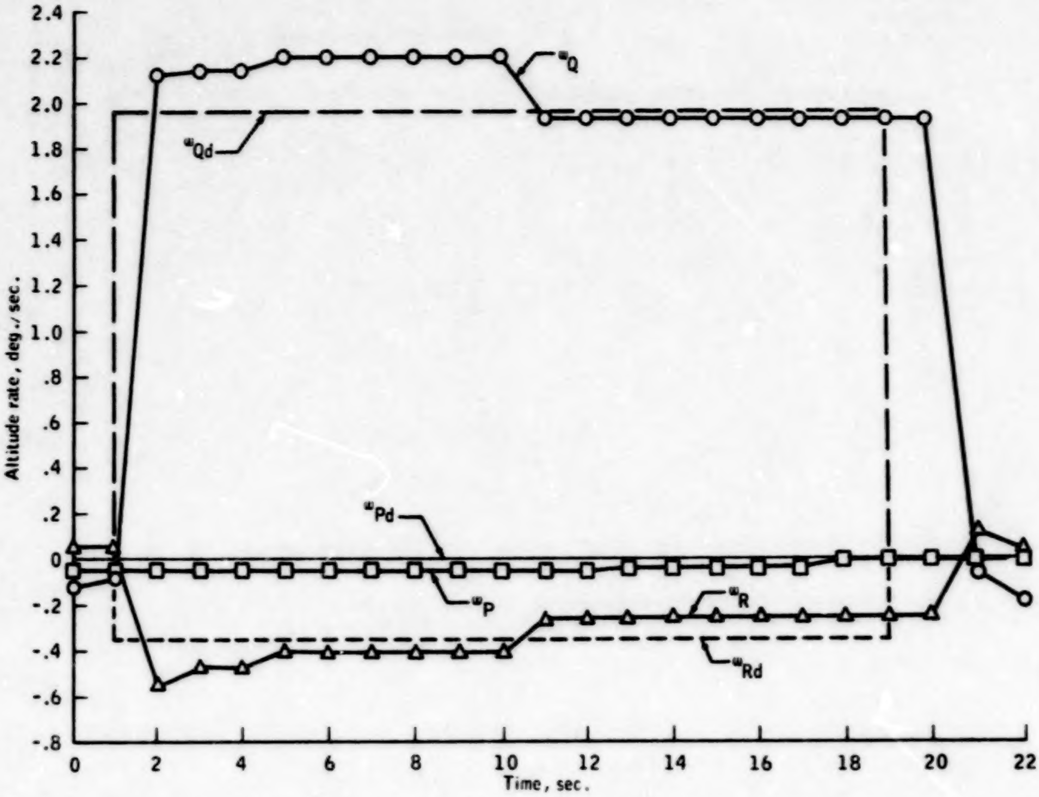


Figure 32. - Attitude maneuver during ascent-coast phase

The final Apollo 9 flight-test results presented (Figure 34) is a plot of phase-plane attitude performance during a DPS insertion firing. The results indicate nominal attitude-hold performance.

CONCLUSIONS

The history of the design and development of a first-generation digital control system has been presented in this case study. Because of the design flexibility inherent in digital systems, it is expected that increased emphasis will be placed upon digital control-system techniques.

This case study has been written to define the environment in which control-system designer works. Reviewing the design history, one gains an implicit perspective viewpoint of the relative importance of optimal control and modern control techniques.

Some of the experience learned during development of the DAP may be used to avoid future design problems. Logical decision techniques should be applied with care in the design development, because conditions may exist in which these techniques may un-

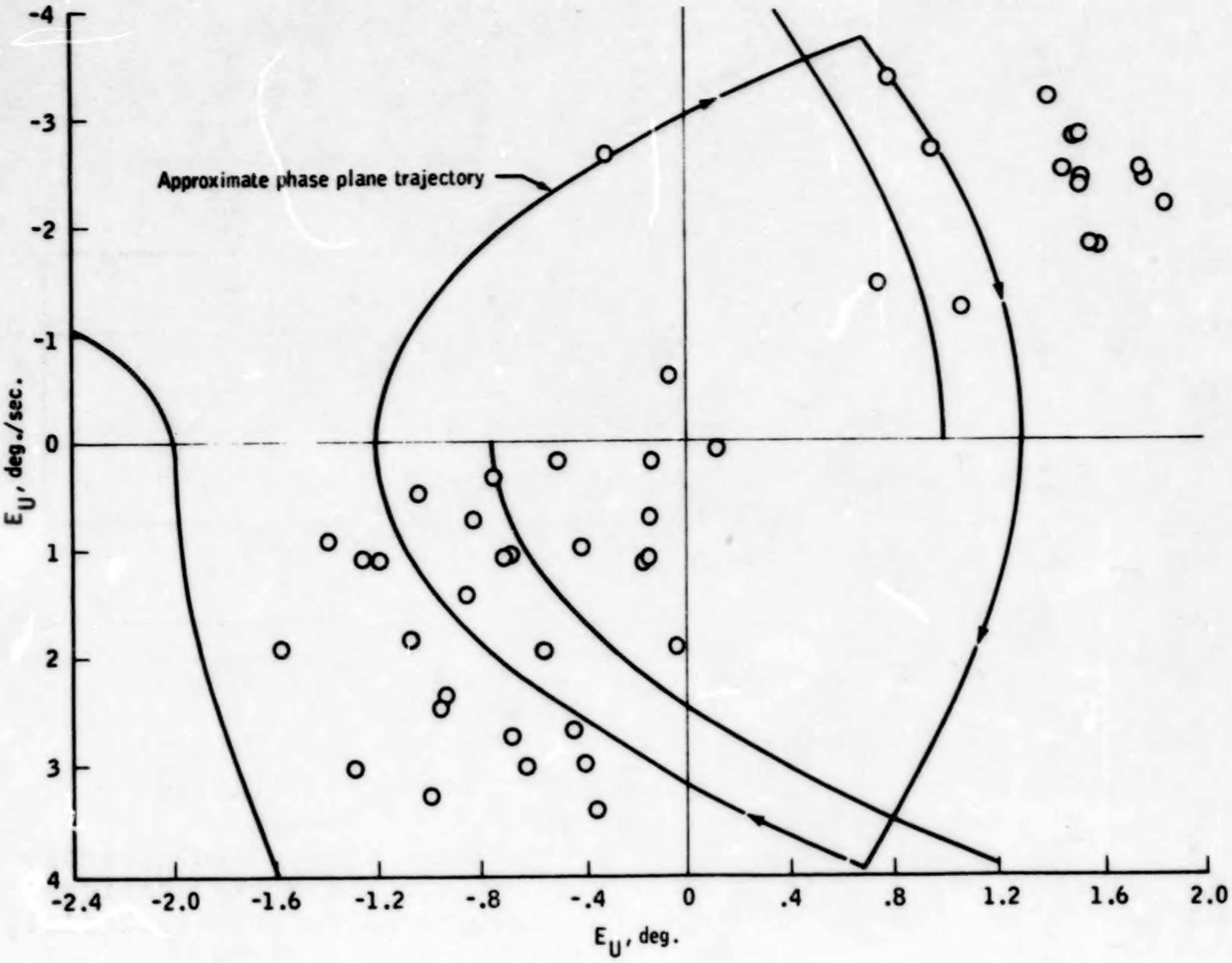


Figure 33. - Limit cycle during LM ascent firing

expectedly lock out entire system functions. The use of logic in avoiding degraded performance has to be traded off with potential unintended restrictions.

Another generalization concerns the manner in which requirements in the estimation function are established. Open-loop testing alone is not always adequate to assess the acceptability of the filter performance. Estimation requirements should reflect the manner in which the output information is used in the control law. As an example, a control law that is mechanized to operate on the sign of a function only has different requirements from a law that operates on both sign and magnitude.

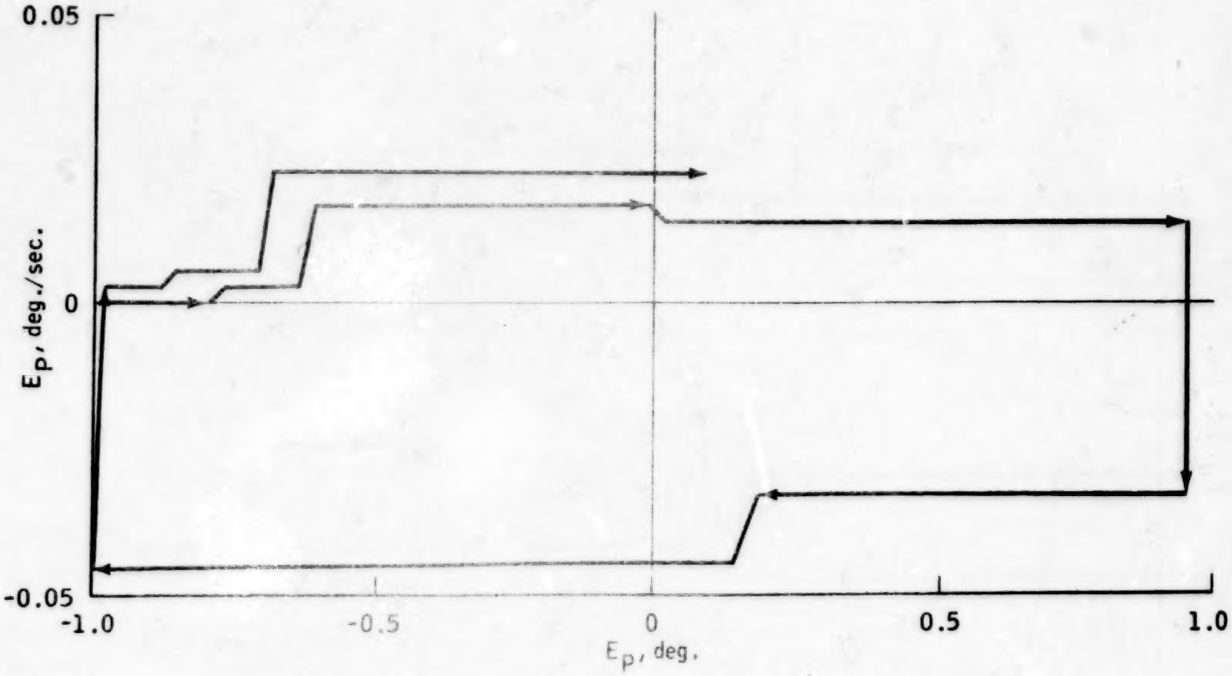


Figure 34. - Limit cycle during LM descent firing

Further research effort should be expended to develop additional analytical techniques for digital control-system design. Adaptive design techniques making use of the inherent flexibility available in digital systems should also be established.

REFERENCES

1. "Guidance System Operations Plan AS-206," MIT Instr. Lab., Rep. R-527, Vol. I, Rev. 1 (Jan. 1967)

2. Covey, R. R., R. B. Sherwood, and S. H. Wu, "Apollo Digital Reaction Control System Design Studies," TRW Space Technology Lab., Rep. 4181-6010-KU000 (Apr. 5, 1965)

3. "Guidance System Operations Plan for Manned LM Earth Orbital Missions Using Program SUNDANCE," MIT Instr. Lab., Rep. R-557, Sec. 3 (Sept. 1968)

4. "Design Verification and Performance Evaluation of the Lunar Module Digital Autopilot," TRW Systems Group, Rep. 05952-6078-R0-00 (Nov. 1966)

5. "Preliminary Analysis of Sampling Rate Requirements and Digital Computer Functional Considerations for Design of LEM Digital Autopilot," GAEC, Rep. LMO-500-211 (Sept. 9, 1964)

6. Heiber, A., J. M. Nervik, and E. A. Nussbaumer, "Apollo Block II CM and LM Guidance and Control Study," Bellcomm, Inc., TM-64-1021-3 (Dec. 17, 1964)

7. Cherry, George, and Joseph O'Connor, "Design Principles of the Lunar Excursion Module Digital Autopilot," MIT Instr. Lab., Rep. R-499 (July 1965)

8. "Performance and Analysis of GFE SUNDANCE Digital Autopilot," GAEC, Rep. LED-500-23 (Feb. 1969)

9. "Evaluation and Analyses of GFE Digital Autopilot (DAP) Performance," GAEC, Rep. LED-500-17 (July 1967)

10. "LM Digital Autopilot Simulation Results Using Program SUNDANCE," MIT Instr. Lab., Rep. E-2377 (Jan. 1969)

11. "Digital Autopilot Software Verification Test Results Summary for Mission D - SUNDANCE and COLOSSUS IA," TRW Space Technology Lab., Rep. 11176-H149-R0-00 (Feb. 1969)

12. "LM-1 Guidance and Control Systems Postflight Analysis," GAEC, Rep. LMO-500-657 (Apr. 5, 1968)

13. "Supplement 4 to Apollo 5 Mission Report - Analysis of Digital Autopilot Performance," TRW Space Technology Lab., Rep. 11176-H061-R0-00 (Oct. 1968). Also available as MSC-PA-R-68-7, Supplement 4 (Jan. 1969)

DESIGN DEVELOPMENT OF THE APOLLO COMMAND AND SERVICE MODULE THRUST VECTOR ATTITUDE CONTROL SYSTEMS

By William H. Peters
Manned Spacecraft Center

SUMMARY AND CONCLUSIONS

Development of the Apollo thrust vector control digital autopilot (TVC DAP) is summarized. This is the control system that provides pitch and yaw attitude control during velocity change maneuvers using the main rocket engine on the Apollo service module.

A list of 10 primary functional requirements for this control system are presented, each being subordinate to a more general requirement appearing earlier on the list. Development process functions are then identified and the essential information flow paths are explored. This provides some visibility into the particular NASA/contractor interface as well as relationships between the many individual activities.

Seven specific problem areas that existed during the development are discussed in detail, including the solutions that evolved. The report concludes with a discussion of flight test data from the first four manned Apollo flights that are pertinent to the performance of this particular attitude control system.

This paper was written primarily for its historical value, as opposed to significances of the particular problem solutions. The primary conclusion regards the benefits inherent in mechanizing controller logic and dynamics in a digital computer. This has provided the flexibility necessary to avoid expensive hardware changes and potential schedule delays. The feasibility and reliability of this approach have been thoroughly demonstrated by the Apollo program.

INTRODUCTION

The purpose of this paper is to record specific design problems that existed during the development of the Apollo thrust vector control digital autopilot (TVC DAP). These subjects are treated, as nearly as possible, in chronological order. To provide the proper framework for this discussion, the entire TVC DAP development process is briefly summarized, resulting in a condensed design history. The control systems that are the subject of this paper have been described in detail elsewhere (references 6 and 7), and the material presented here is descriptive only to the point of illustrating design problems and their solutions.

Since it is frequently convenient to abbreviate the names of complete subsystems by use of the initials from the subsystem name, a list of these acronyms is presented below for easy reference.

LIST OF ABBREVIATIONS

CSM	Apollo command and service module.
CSM/LM	Apollo CSM with lunar module docked.
SPS	Service propulsion system. The rocket engine system on the Apollo service module.
TVC	Thrust vector control. This term is sometimes used to define the SPS gimbal actuation system only, but in the context of this report, it applies to the total problem of orienting the thrust vector to the spatial attitude commanded by the Apollo navigation and guidance systems.
CMC	Command module computer.
DAP	Digital autopilot.
TVC DAP	The particular Apollo DAP used during SPS burns.
SCS	Stabilization and control system. The analog autopilots of the Apollo CSM.
YACTOFF	Yaw actuator offset for trim position.
PACTOFF	Pitch actuator offset for trim position.
RO3	Software routine for DAP initialization.
GSOP	Guidance systems operation plan.
NASA	National Aeronautics and Space Administration.
MSC	Manned Spacecraft Center.

REQUIREMENTS

It is usually difficult to list a comprehensive and realistic set of requirements for any complex subsystem at the onset of a development program. This is primarily due to an initial lack of knowledge concerning all the interactions between the particular subsystem and the total system environment with which the subsystem must ultimately be compatible. However, the multitude of "requirements" that gradually evolve during the development process are primarily design details that can be considered subordinate to a set of more basic functional requirements. The hierachy of functional requirements for the Apollo CSM powered flight attitude control system were as follows:

1. Maintain attitude orientation sufficiently near a commanded orientation during accumulation of velocity from a service propulsion system (SPS) rocket engine burn;
2. Provide acceptable thrust impulse economy;
3. Provide an attitude control loop responsive enough to prevent dynamic interference with the guidance scheme;
4. Provide compensation for rigid body motion, including time varying gains;
5. Provide compensation for initial thrust vector misalignment;
6. Provide compensation for fuel sloshing and structural bending resonances;
7. Provide stability margins of sufficient magnitude to guarantee stability of (a) guidance long period mode; (b) attitude short period mode; (c) fuel sloshing resonances; and (d) structural bending resonances.
8. Provide acceptable performance in convergence from large transients and in the steady-state limit cycles;
9. Maintain attitude errors and vehicle rates within bounds compatible with effective crew monitoring of successful performance; and
10. Provide a capability for switching to a redundant controller in case of malfunction.

The terms "sufficiently near," "acceptable," and the like, were not specified quantitatively in all cases. However, they were closely monitored by cognizant NASA engineers during the development phase and requirements for design improvements were negotiated on several occasions.

THE DEVELOPMENT PROCESS

The significant elements of the development process are diagrammed in Figure 1. The discussion below elaborates on some of the blocks in the diagram. Note that much of the activity revolved around the block labeled "modifications." The flexibility for incorporating design modifications into a digital control system,

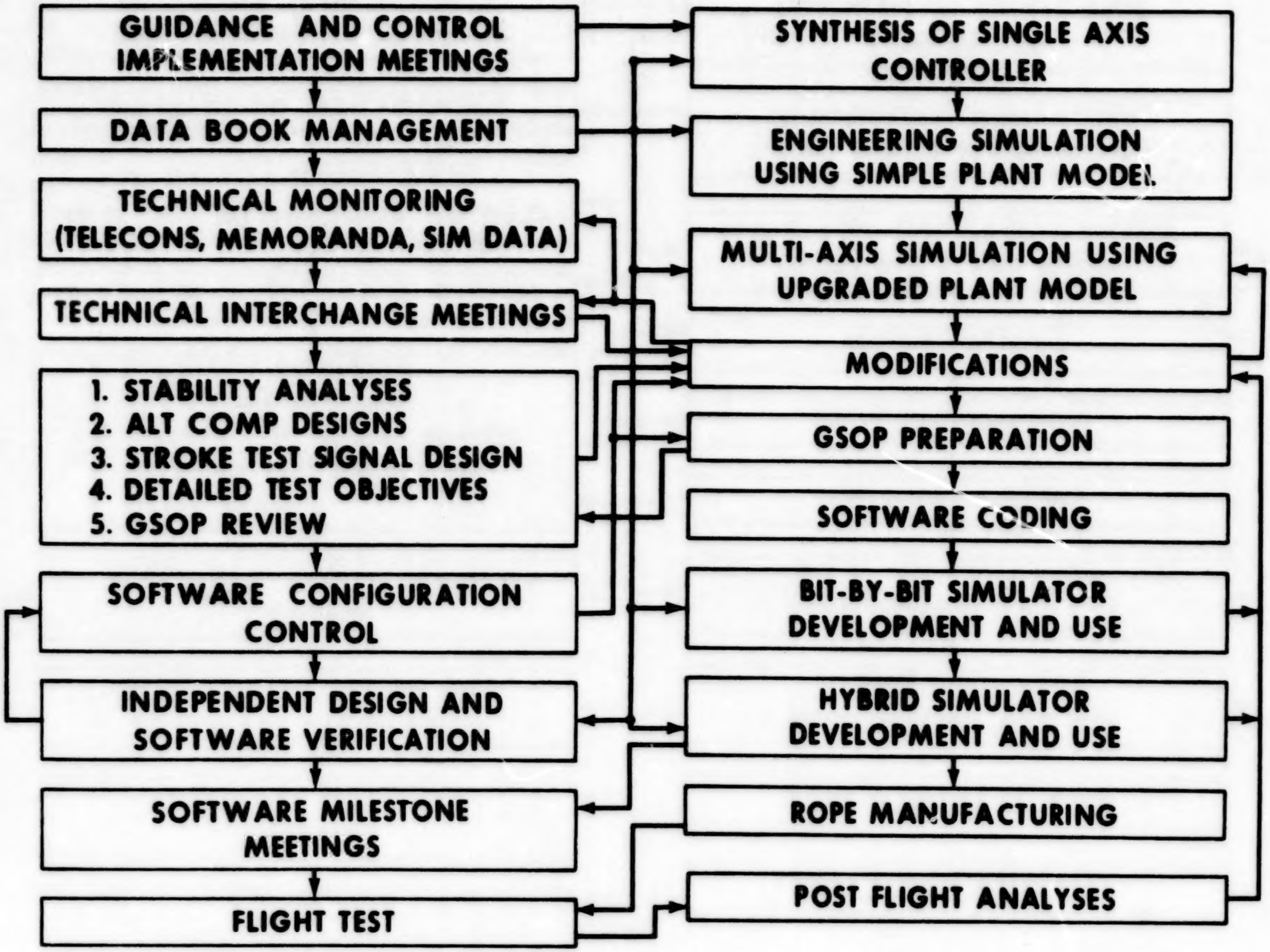

GUIDANCE AND CONTROL IMPLEMENTATION MEETINGS
DATA BOOK MANAGEMENT
TECHNICAL MONITORING (TELECONS, MEMORANDA, SIM DATA)
TECHNICAL INTERCHANGE MEETINGS
1. STABILITY ANALYSES
2. ALT COMP DESIGNS
3. STROKE TEST SIGNAL DESIGN
4. DETAILED TEST OBJECTIVES
5. GSOP REVIEW
SOFTWARE CONFIGURATION CONTROL
INDEPENDENT DESIGN AND SOFTWARE VERIFICATION
SOFTWARE MILESTONE MEETINGS
FLIGHT TEST
SYNTHESIS OF SINGLE AXIS CONTROLLER
ENGINEERING SIMULATION USING SIMPLE PLANT MODEL
MULTI-AXIS SIMULATION USING UPGRADED PLANT MODEL
MODIFICATIONS
GSOP PREPARATION
SOFTWARE CODING
BIT-BY-BIT SIMULATOR DEVELOPMENT AND USE
HYBRID SIMULATOR DEVELOPMENT AND USE
ROPE MANUFACTURING
POST FLIGHT ANALYSES

Figure 1.- TVC DAP development process

in a straightforward, easy, and inexpensive manner (compared to changes in conventional hardware systems), and also much later in the program than would normally be allowed, gives the designer much more freedom both in development time schedule and the type of changes that will be tolerated.

The activities diagrammed in Figure 1 roughly depict the NASA functions on the left and the contractor functions on the right. It should be kept in mind, however, that no such "lumped" representation can be exact, and that the flow lines only approximately represent the interactions that actually occurred.

The technical interchange meetings proved to be an invaluable extension of the technical monitoring function for keeping track of the development status and planning for future emphasis on the current problems (specific problem areas are discussed in the next section). The GSOP (guidance system operation plan) constituted the primary control documentation, and the GSOP review function provided the checkpoint for assuring that detailed requirements would be implemented.

The bit-by-bit simulations were digital programs that caused a general-purpose computer to function like the airborne computer hardware, and provide detailed diagnostic information flight software coding problems. The hybrid simulations coupled actual flight-type hardware into computer simulations of spacecraft dynamics. The independent design and software verification consisted primarily of simulations performed at NASA/MSC and the spacecraft contractors, where both bit-by-bit and hybrid simulations were also used.

After the software had developed to the point where the schedule required that configuration control be imposed, the software milestone meetings gradually replaced the need for the technical interchange meetings. Beyond this point, improvement changes were allowed only if justifiable to the software configuration control board.

DEVELOPMENT PROBLEMS AND SOLUTIONS

This section consists of several, somewhat self-contained, discussions of the major areas that caused concern. Included are statements of the solutions that evolved.

CSM/LM Passband Requirement

Due to the fact that the guidance scheme to be used had time-varying gains, and the fact that the higher the guidance loop gains became the more rapidly they changed, it was originally presumed that a linearized stability analysis of the combined guidance and attitude control dynamics was not too meaningful. This approach showed that the system would go unstable for some value of "time-to-go" (T_{GO}), regardless of the control system passband (frequency beyond which sinusoidal attitude command components would be attenuated and delayed).

The early thoughts on this subject were usually resolved by the argument that T_{GO} computations would have to be frozen at some point, and only simulations could realistically resolve the question of whether or not the control system was sufficiently sluggish to impair the guidance performance.

The passband was constrained on the upper side by slosh and bending stability considerations. Lack of good quality bending data (discussed in more detail as a separate item) forced the early design effort to assume that the passband requirement was "as low as possible and not interfere with guidance." This was generally assumed by the G&C contractor to be somewhere between 0.1 and 0.5 rad/sec. The NASA position was that the gains should be as high as possible, consistent with adequate high-frequency stability margin and large transient stability. Hence, due primarily to a lack of data on the flexible characteristics of the vehicle, and a lack of experience with the particular guidance law to be used, the program was denied an early specification of this most basic servo design parameter.

Work performed by the G&C contractor showed that lunar orbit insertion errors began to diverge when the guidance loop gain (deg/sec vehicle rotation rate command per degree of vehicle mass acceleration error) was reduced below 0.1 per second, but these studies excluded attitude control dynamics (ref. 1). On the assumption that 0.1 was an acceptable guidance gain, and that control loop dynamics would be acceptable as long as a linearized model of guidance and control dynamics were stable for $T_{GO} = 20$ seconds, the control-loop passband requirement was established by the Guidance and Control Division at the MSC. A second-order differential equation model of the attitude control dynamics was coupled to a linearized model of the powered flight guidance law for T_{GO} frozen at 20 seconds. The natural frequency of the control quadratic was then lowered to the onset of instability. This was found to be in the neighborhood of 0.4 rad/sec, which was used as the baseline design requirement until the summer of 1966.

SPS Actuator Performance Degradation

During the summer of 1965, the prime contractor for the Apollo command and service module (CSM) learned that their subcontractor for the SPS engine actuator was having difficulty in meeting performance specifications. The electro-mechanical (motor, gears, and jack screw) actuator was supposed to be able to slew the SPS engine position at a steady-state rate of 0.36 rad/sec, but first-delivered units were only capable of 0.2 rad/sec. Severe heating problems created a requirement to reduce the maximum actuator rate even further or subject the actuator to major redesign. This caused NASA and their contractors to try and find the minimum engine actuator rate that would be compatible with Apollo attitude control requirements.

The factors necessary to establish this requirement were as follows:

1. Attitude control passband requirement (set by guidance requirement);
2. Flexible body stability (in order to guarantee stability margin by an analytical process, it is necessary to avoid significant non-linearities);
3. Rigid body recovery from large transients;
4. Cost and schedule; and
5. Sound system design principles.

Items 1 and 2 above established only mild requirements for fast actuator response. If it is assumed that the control system passband requirement is 0.4 rad/sec (the minimum requirement established in the previous section), and further, on the assumption that the vehicle must respond to a sinusoidally varying rate command from guidance with rates as high as 0.35 rad/sec before saturating the thrust actuator (which simply means that a linear math model breaks down at this amplitude), the following approximate, minimum actuator rate requirement can be derived:

$$\dot{\theta} \simeq \dot{\theta}_c = 0.35 \sin (0.4 t) \tag{1}$$

after two differentiations

$$\dddot{\theta} \simeq -(0.4)^2 (0.35) \cos (0.4 t) \simeq \mu_c \dot{\delta} \tag{2}$$

$$\dot{\delta} = \frac{0.056}{\mu_c} \text{ rad/sec} \tag{3}$$

where

$\dot{\theta}$ = actual vehicle rate in response to a sinusoidal attitude rate command from guidance

$\dot{\theta}_c$ = guidance command

μ_c = vehicle angular acceleration coefficient

$\dot{\delta}$ = rate of gimballing rocket engine relative to vehicle airframe.

NOTE: The approximately equal signs ($\simeq$) are required because some dynamics are ignored (e.g., $\dot{\theta}$ may be slightly greater than $\dot{\theta}_c$ if the attitude control system is underdamped).

For Apollo, the smallest value of μ_c was approximately unity. Hence, the maximum actuator rate required to assure non-saturation due to guidance alone would be less than 0.1 rad/sec.

The requirement for linear response to structural oscillations is more difficult to obtain, but can be estimated by assuming that only one or two of the low order modes are excited to near the point of structural failure. This is then interpreted in terms of actuator command as a function of where the attitude sensor is located, and the dynamic gain in the controller. Due to the extreme attenuation being proposed for the CSM/LM TVC DAP at frequencies near and above the first mode resonance, the actuator command was essentially independent of the bending motions.

Item 3 above proved to be the major constraint on the SPS actuator rate requirement. For the low actuator rate limits being considered (0.1 rps) and the backup analog control system (SCS) design that existed at that time (mid 1965), vehicle attitude would not converge properly from large transient conditions (refs. 2,3). Fortunately, the actuator rate limit did not severely constrain the DAP design, but the degraded actuator performance did require analysis and simulation emphasis on the large signal performance.

Item 5, simply stated, means that a designer must not overconstrain the design problem, which implies conservatism when constraints are being adjusted, because it is difficult to predict all the possible ramifications. However, in this case the cost and schedule implications (payoff function) were heavily weighted toward accepting some risk of later difficulty in other areas (e.g., perhaps creating a requirement for future changes to hardware in the control system). Due to these conflicting factors, the minimum specification finally agreed upon between NASA and the spacecraft contractor (0.1 rad/sec) was not conservative; however, it was generally expected by both that actual delivered hardware would be considerably better than this minimum requirement. The final hardware was capable of slewing the SPS engine at about 0.15 rps, and SCS gains had been lowered such that adequate large-signal stability was obtained.

Digital Filter Mechanization

The single-axis attitude control loop compensation for Lunar Module attached was synthesized by the G&C contractor to be a seventh-order filter. The first attempt at a digital mechanization of these dynamics was a straightforward, recursive computation, performed by shifting all data by one step each computation cycle before multiplying by the appropriate coefficients and summing. The process is shown schematically in Figure 2, where Z^{-m} represents a delay of m computational intervals. The gains are simply the coefficients of the numerator and denominator polynomials of the Z-transform version of the filter dynamics (ref. 4). The reference refers to this method as "direct programming."

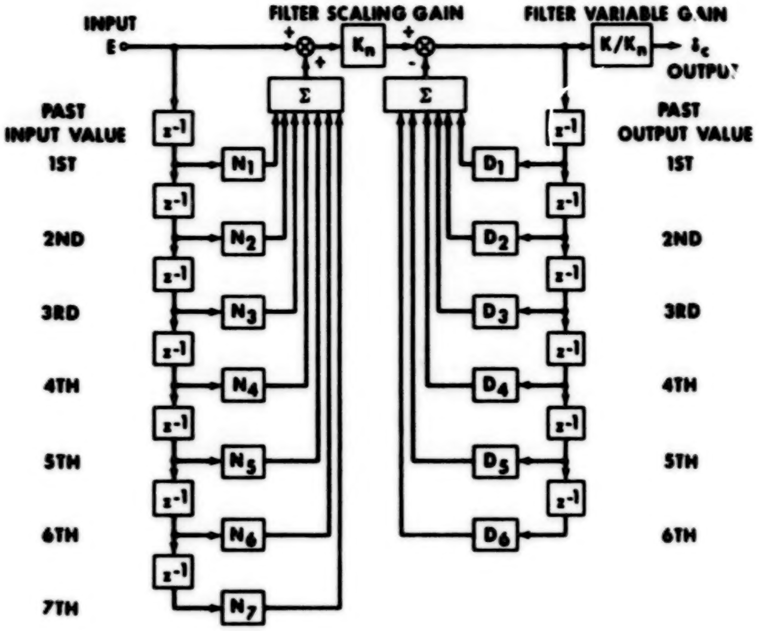

INPUT
E
FILTER SCALING GAIN
+
+
K_n
+
-
FILTER VARIABLE GAIN
K/K_n
δ_c
OUTPU
Σ
Σ
PAST
INPUT VALUE
1ST
2ND
3RD
4TH
5TH
6TH
7TH
z^{-1}
N_1
N_2
N_3
N_4
N_5
N_6
N_7
D_1
D_2
D_3
D_4
D_5
D_6
PAST
OUTPUT VALUE
1ST
2ND
3RD
4TH
5TH
6TH

Figure 2.- CSM/LM filter

This approach did not work satisfactorily, and the reasons are discussed in detail in reference 5. Briefly, the cause was due to repetitive truncation error in the computational process, because each intermediate result had to be stored in a finite word-length register. The effect was to shift the effective location of some filter poles into the right-half plane causing the filter alone to be unstable. After some study of this problem by the G&C contractor, the second approach discussed by reference 4, called "iterative programming," was adopted. Essentially, this technique separates the factored form of the filter transfer function into a product of two or more (three in the case under discussion) smaller filters, implements these, and lets the output of the first be the input to the second, and so forth. The final Apollo TVC DAP mechanization is discussed in references 6 and 7, and the earlier version is discussed in reference 8.

Plant Model Deficiencies

A straightforward synthesis of required filter dynamics would have been possible only if there had been complete confidence in the mathematical model of the "plant" dynamics. This includes not only the mathematical formulation of the model, but also the computer programming of that model into the simulations being used to verify the filter design. Generally, the required mathematical formulations have existed for some time, but a problem exists due to the multitude of variables and a desire to eliminate unnecessary coupling terms to reduce simulator complexity. This usually results in reduced equation sets being published in various reports, and it is impossible to determine if a specific reduced set is applicable to a specified situation unless the missing terms are available for inspection and evaluation. Two examples of this type of problem that arose during the development of the Apollo CSM TVC DAP are cited below.

Early evaluation of the single-plane dynamics of the Apollo CSM/LM, performed independently by NASA/MSC and a support contractor consistently disagreed with the G&C contractor results regarding stability of propellant slosh resonances. The G&C contractor results showed the CSM propellant sloshing to be most severe, but with acceptable stability margin, whereas the other analyses indicated the LM propellant sloshing to be most severe and to have unacceptable stability margins. A thorough investigation of the models showed that coupling with the translational degree of freedom produced this disparity of results. The MSC and support contractor models had both been derived under the assumption that this coupling was small enough to be ignored for evaluation of "high-frequency" rotational dynamics.

NASA/MSC studies (ref. 9) had shown that there was a possibility of the combined airframe/actuator dynamics being unstable without any attitude controller inputs to the actuator. The energy source was the thrust force, and the feedback path was inertial forces on the engine, due to structural oscillations, causing engine gimballing motion. This effect is sometimes referred to as "dog-wags-tail" as opposed to the other situation known as "tail-wags-dog." Early contractor models had omitted these terms (ref. 10). This coupling does reduce damping of the first two structural resonances of the Apollo CSM/LM vehicle, but not to the point of instability. If it had, then the adopted DAP filter design approach of providing large attenuation to signals due to bending (gain stabilization) would not have produced an acceptable design.

Lack of Good Quality Bending Data

The former section dealt with model formulation and simulator programming accuracy, whereas this section deals with the availability of data to use in the models. This was the singularly large, overriding problem of the entire TVC DAP development process. The troubles began in the spring of 1965 when it was decided that a combined CSM/LM structural dynamics test, which had been planned, could be eliminated from the program. This was later compounded when it was decided that the in-line responsibility for combined vehicle dynamics would be at MSC instead of with one of the Apollo spacecraft prime contractors. The result was considerable confusion regarding data to define airframe dynamics.

The idea of an in-flight dynamics test was proposed by the Guidance and Control Division in October 1965 (ref. 11). The structural dynamics group had hopes of getting some data from this test that could be used to calibrate their analytical models, but the required additional telemetry measurements were not sanctioned by program management. A limited version of the originally planned test was performed and is discussed below under Flight Test Results.

An MSC contractor computed the first set of three-dimensional mode shapes for the combined Apollo vehicle in February 1966, two full years after the initial TVC controller synthesis. This effort predicted a first-mode frequency about one octave lower than former predictions (roughly 1.0 Hz as opposed to 2.0 Hz). Obviously this seriously degraded the confidence in the quality of any bending data in existence at that time, and probably was the most significant factor in getting the dynamics ground test of the combined vehicle back into the program.

The ground test was reinstated, but due to scheduling problems, test preparation, test performance, data reduction, and so forth, the results were not available until December 1968. In the interim the detailed structural dynamic modeling being performed by NASA produced the first set of realistic, three-dimensional mode shapes, part of which were made available in June 1967. This essential lack of credible bending data for the 18-month period encompassing 1966 and the first half of 1967 precipitated the next two problems discussed below.

Mod 40 to Mod 80

During the summer of 1966, just prior to publication of the GSOP that would specify the software to be used in the first mission containing a docked CSM/LM SPS engine burn (ref. 7), the G&C contractor decided that additional conservatism regarding high-frequency stability margin was required. (Recall that the estimates of first bending mode frequency existing at that time was nearly a full octave below the nominal first mode frequency used in the initial DAP filter synthesis.) The design change that was implemented at this point is a good example of the flexibility available to the digital control system designer that would be impossible if the control system design were being translated into hardware.

Briefly, the change represented a radical departure from the former concept of a single-stage compensation filter with fixed dynamics and time-varying gain. The original design had the filter coefficients selected for a sample interval of once each 40 milliseconds (Mod 40). The new concept was to start the burn with the originally designed filter, but to switch gains and sampling interval after a partial nulling of the start transients. A thrust misalignment correcting inner loop was also added (discussed in next section).

The post-switch filter dynamics, due to a sampling interval of 80 milliseconds (Mod 80), was essentially the same as that of the Mod 40, except that the real frequency where a given shaping effect would occur was not one full octave lower. This characteristic of a digital filter can be explained by reference to the relationship between real frequency w and the fictitious W-plane frequency v.

$$v = \tan \frac{wT}{2}$$

where

v = W-plane frequency

w = real frequency

T = sample period.

Note that a specific value of v, say v_o, is uniquely related to real frequency only through the sampling interval. Hence,

$$v_o = \tan \frac{w_o T_o}{2} = \tan \left[\frac{1}{2} (w_o/K) \; (KT_o) \right] \tag{5}$$

implies that the specific real frequency, where a given amount of attenuation and phase shift would occur, is shifted by the inverse of any scale factor applied to the sampling interval. This is obviously an extremely important flexibility available in a digital filter that is not available in filters implemented by hardware components. Hardware filters have an inherent sense of real time, based upon the physical and geometrical aspects of the hardware, whereas a digital filter is a mathematical entity, and hence insensitive to real time. It is for this reason that extremely low frequency filtering can be done digitally in situations where it would be impossible to do the same job with hardware.

Mathematical proof that real-frequency scaling and sample-period scaling are truly related as implied above is beyond the scope of this report. However, a rather convincing argument is readily available. Since the digital filter is synthesized in the W-plane (poles and zeros arranged to supply a specific filtering action for a specific value of W-plane "frequency" v_o), and since the bilinear transformation into the Z-plane is not a function of sampling interval, then a given set of digital filter coefficients will provide the same filtering for a given v_o, regardless of the sampling interval involved.

This particular change is listed as one of the TVC DAP development problems, because the passband of the Mod 40 design (about 0.6 rad/sec) was already approaching the minimum value, and this change brought it down to around 0.3 rad/sec. The resulting performance was quite poor, requiring a considerable amount of effort in "tuning up" the guidance and control interface, and trying to contain the attitude errors within reasonable bounds. The result was NASA direction to the contractor (ref. 12) to produce an uprated TVC design, based on the assumption that the TRW bending data frequencies were accurate within ±30 percent (refs. 13, 14).

Guidance and Control Interaction

The low gain attitude control loop would permit extremely large peak attitude error transients and steady-state offset attitudes, resulting from initial thrust torque bias (engine misalignment). The design change implemented to attack this problem, still assuming that the low gain was necessary for conservatism on high-frequency stability margin, was to add an inner loop that would extract the average commanded engine position and feed this back to the engine position command as a bias. The digital filter used for this purpose had large time constants to avoid interaction with the attitude short period mode, resulting in additional phase lags at frequencies which now coupled with guidance (outer loop) performance. Also, the long time required to initialize properly the thrust position filter (find the trim position) reduced the effectiveness of this approach for limiting the initial attitude excursions. Hence, large guidance errors could still accumulate during the first few seconds of a burn which could not be completely removed in short burns. (An example of this is discussed in a following section on flight test results.)

The above does not mean to imply that the thrust misalignment correction loop was completely ineffective. Its main purpose was to remove large steady-state attitude errors required to provide acceleration trim, so that attitude error could be used by the pilot to determine if the maneuver were progressing normally.

Mechanization of the guidance/control interface can produce, or aggravate, guidance and control interaction problems. The evolution of the interface used in the Apollo guidance and control system is shown in figures 3 through 5. The guidance law produces a desired acceleration vector rate command that is proportional to the sensed acceleration direction error. Since the steady-state acceleration direction is very nearly the attitude of the vehicle longitudinal axis (this is not true when control torques are being commanded), there is a requirement to provide stable attitude inner loop control by using attitude position feedback.

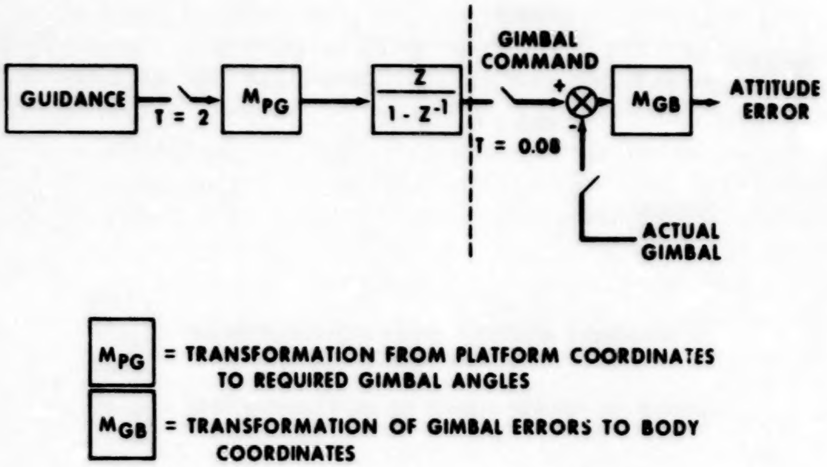

GUIDANCE
T = 2
M_PG
Z / 1 - Z^-1
GIMBAL COMMAND
+
-
T = 0.08
M_GB
ATTITUDE ERROR
ACTUAL GIMBAL
M_PG = TRANSFORMATION FROM PLATFORM COORDINATES TO REQUIRED GIMBAL ANGLES
M_GB = TRANSFORMATION OF GIMBAL ERRORS TO BODY COORDINATES

Figure 3.- Original guidance and control interface

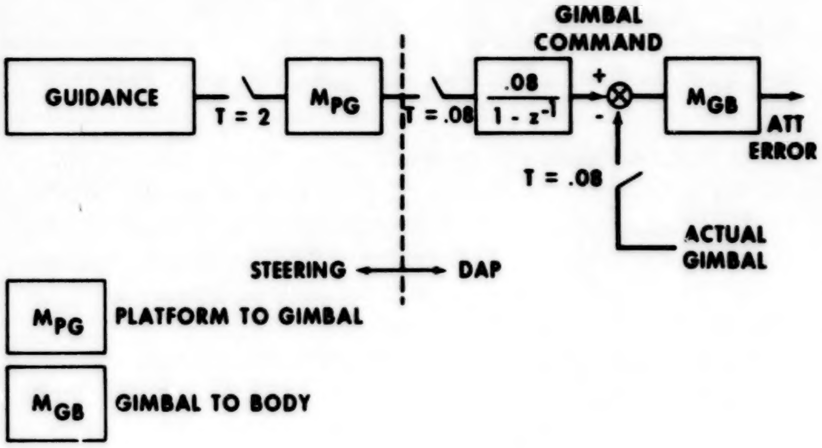

GUIDANCE
T = 2
M_PG
T = .08
.08 / 1 - z^-1
GIMBAL COMMAND
+
-
M_GB
ATT ERROR
T = .08
ACTUAL GIMBAL
STEERING
DAP
M_PG PLATFORM TO GIMBAL
M_GB GIMBAL TO BODY

Figure 4.- An alternate guidance and control interface (ramp steering)

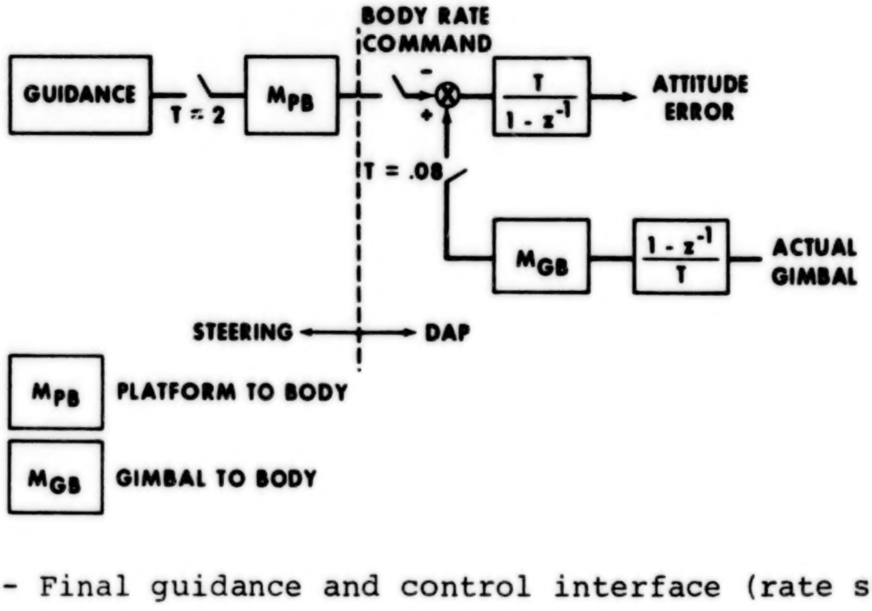

GUIDANCE
T = 2
M_PB
BODY RATE COMMAND
-
+
T / 1 - z^-1
ATTITUDE ERROR
T = .08
M_GB
1 - z^-1 / T
ACTUAL GIMBAL
STEERING
DAP
M_PB PLATFORM TO BODY
M_GB GIMBAL TO BODY

Figure 5.- Final guidance and control interface (rate steering)

The first interface mechanization performed a simple digital integration of the guidance rate command (see Figure 3) to form a vehicle attitude command. This caused step changes in the attitude command each guidance computation period (2.0 seconds) which caused significant interaction with propellant slosh dynamics. The next approach was to smooth the attitude commands by performing the integration at the DAP sample rate (see Figure 4). This change was effective in reducing the slosh excitation, but there was some concern about the fact that gimbal angle errors were being subjected to a coordinate transformation. This cast some doubt on large signal stability. Subsequently the coordinate transformation problem was solved, as shown in Figure 5, by transforming guidance rate command to body coordinates where it is compared with a derived body rate before the integration to produce body attitude error.

FLIGHT TEST RESULTS

Apollo 7 was the first manned Apollo flight and the first to use a digital autopilot. This flight consisted of the Apollo CSM only (a different digital filter is required when the Lunar Module is attached). The flight contained eight main engine burns, which are summarized below.

There were five burns under TVC DAP control for the entire burn, and a sixth that was initiated by the primary control system with a switch to a backup manual control mode after 35.8 seconds. Two were minimum impulse burns and three burns ranged between 7.8 and 10.9 seconds duration. Thrust mistrim at the beginning of these burns was less than 0.2 degree in all cases, and peak attitude errors (except roll) were all less than 0.5 degree during DAP control.

Short tables are presented to summarize the significant flight test results for the TVC DAP from Apollo 8 through 10. These are Tables I through III respectively. Data from Apollo 7 are omitted because they were similar to Apollo 8, and data from Apollo 11 are omitted because they were similar to Apollo 10.

The Apollo 8 lunar flight used essentially the same control systems that were used by Apollo 7 with equally good results. This provided additional flight experience, especially in the area of long-duration burns (Table I). Control system dynamics interaction with guidance performance was insignificant for the Apollo CSM only configuration. This was not the case, however, for the CSM/LM configuration, as evidenced from the flight of Apollo 9.

Apollo 9 was the big developmental flight for the TVC DAP design. From the previous sections of this report it should be obvious that the control-loop performance had been degraded to the

TABLE I - APOLLO 8 FLIGHT TEST RESULTS

SPS Burn No.	Vehicle Config.	Total ΔV (fps)	Peak Pitch or Yaw Error (deg)	ΔV Residuals before Nulling		
				$\dot{X}$(fps)	$\dot{Y}$(fps)	$\dot{Z}$(fps)
1	CSM	24.8	0.4944	4.4	-0.1	+0.1
2	CSM	3000.0	0.4504	-1.4	0.0	+0.2
3	CSM	134.8	0.1634	No data	No data	No data
4	CSM	3520.0	0.3955	-0.5	+0.4	-0.1

TABLE II - APOLLO 9 FLIGHT TEST RESULTS

SPS Burn No.	Vehicle Config.	Total ΔV (fps)	Peak Pitch or Yaw Error (deg)	ΔV Residuals before Nulling		
				$\dot{X}$(fps)	$\dot{Y}$(fps)	$\dot{Z}$(fps)
1	CSM/LM	36.8	0.48	1.6	0.5	-0.2
2	CSM/LM	850.6	2.31	0.0	1.0	0.2
3	CSM/LM	2570.7	4.36	2.7	-2.5	-2.3
4	CSM/LM	299.1	2.77	0.2	3.9	2.7
5	CSM/LM	571.8	7.21	1.9	11.4	1.7
6	CSM	38.8	0.41	1.1	-0.6	-0.3
7	CSM	653.3	1.85	-1.3	1.0	-0.2
8	CSM	321.4	1.35	7.5	0.63	-2.0

TABLE III - APOLLO 10 FLIGHT TEST RESULTS

SPS Burn No.	Vehicle Config.	Total ΔV (fps)	Peak Pitch or Yaw Error (deg)	ΔV Residuals before Nulling		
				$\dot{X}$(fps)	$\dot{Y}$(fps)	$\dot{Z}$(fps)
1	CSM/LM	19.7	<0.5	1.0	0.3	0.7
2	CSM/LM	48.7	<0.5	-0.9	-0.1	0.3
3	CSM/LM	2982.4	0.3	0.0	-0.2	0.0
4	CSM/LM	138.9	0.39	0.5	-0.4	-0.4
5	CSM	3630.3	0.60	0.3	1.6	-0.1

point of potentially significant guidance and control interaction, primarily as a safety measure for this first CSM/LM docked flight to guard against high-frequency instability. Also, an in-flight dynamics test had been designed to determine experimentally the frequency response of the airframe dynamically coupled to the SPS engine actuation system (ref. 15). This test consisted of a special routine in the CMC acting as a function generator to supply excitation to the SPS engine pitch command during powered flight. Telemetered response data were then analyzed to determine the total plant transfer function response to verify dynamic stability margin, and provide confidence in the analytical math models. The post-flight analysis of these data appears in reference 16. The essential result of this test was an experimental verification of the coupled plant transfer function, response in the flight environment, in a frequency band encompassing the first two predominant structural resonances. The predicted amplitude ratios at the first two frequencies were -1.5 and -6.6 dB, while those determined from the flight data were -6.0 and -12.0 dB, respectively.

The fifth SPS burn on Apollo 9 exhibited a rather serious deviation of the spacecraft attitude. Data for this burn are shown in Figure 6. Note that the attitude error was quite large throughout the burn, peaking at just over 7.0 degrees. Analysis of these data showed that this type of response was proper for the TVC controller being used on this mission. The main disturbance inputs here were center-of-mass motion (a jerk disturbance) as well as a 0.2-degree initial thrust misalignment (acceleration disturbance). In view of

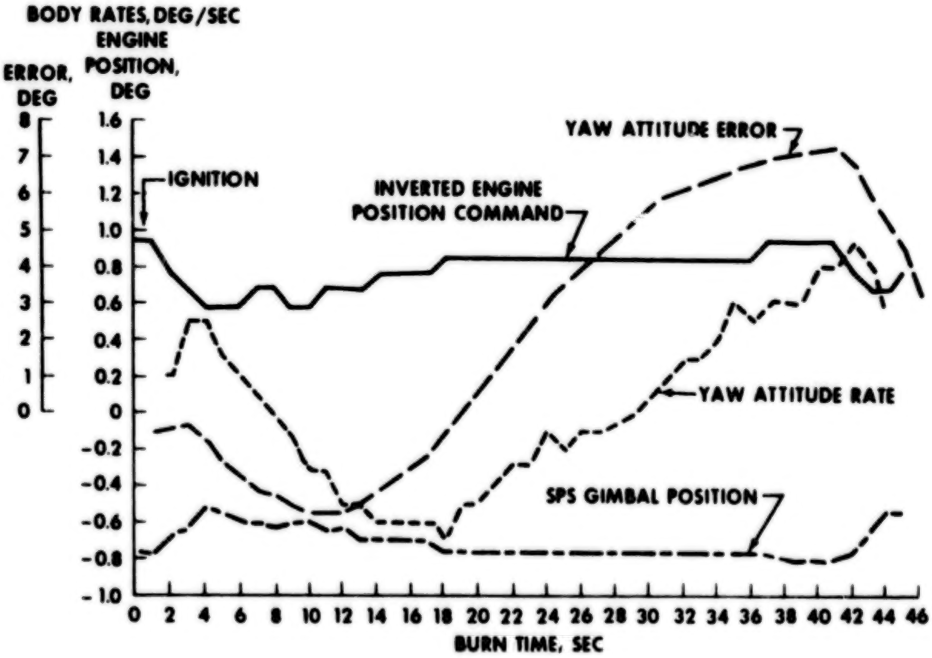


Figure 6.- Apollo 9 yaw data for SPS 5 (data plotted for one second intervals)

the fact that even larger initial thrust misalignments can occur, this performance is considered unacceptable.

Table II provides Apollo 9 performance data for the five CSM/LM burns in the form of peak attitude errors, peak body rates, and residual velocity error at the end of the burns. Note that the velocity error at the end of SPS burn number 5 was 11.4 ft/sec, demonstrating the severity of the control dynamics interacting with guidance performance.

The uprated CSM/LM TVC DAP design had been accepted for incorporation into the flight software that was used on the first lunar mission of the complete Apollo vehicle. As a result, guidance performance was greatly improved. The Apollo 10 performance is summarized in Table III.

The following versions of the Apollo TVC DAP all flew successfully without any known design or programming errors.

1. CSM alone original;
2. CSM alone final (same functional design as 1 but completely different coding);
3. CSM/LM original;
4. CSM/LM final (functional change as well as coding changes).

This apparently speaks well of the basic approach of implementing controller dynamics in a digital computer.

REFERENCES

1. Mertz, A. W.: "Lunar Deboost Digital Computer Studies." Mass. Instit. of Tech/Inst. Lab report SGA Memorandum No. 18-65, 1965.

2. Peters, W. H.: "Effects of Limiters in CSM SPS Actuation System." NASA MSC memorandum EG23-8-66, January 18, 1966.

3. Peters, W. H.: "Block II Actuator Rate Requirements." NASA MSC memorandum EG23-66-51, May 10, 1966.

4. Tou, J. T.: Digital and Sampled Data Control Systems. McGraw-Hill Book Company, Inc., New York, N. Y., 1959.

5. Whitman, C. L.: "The Implementation of Digital Filters in Computers of Small Word Length." Mass. Instit. of Tech./Thesis for Master of Science Degree, February 1966.

6. "Guidance System Operations Plan for Manned CM Earth Orbital and Lunar Missions Using Program COLOSSUS 2C "(Commanche Rev. 67), Section 3 Digital Autopilots (Rev. 7); Mass. Instit. of Tech./ Inst. Lab report R-577, August 1969.

7. Stubbs, G. S.; Penchuk, A.; and Schlundt, R. W.: "A Digital Autopilot for Thrust Vector Control of the Apollo CSM and CSM/LM Vehicles." Mass. Instit. of Tech./Inst. Lab report, March 17, 1969.

8. "Guidance System Operations Plan AS 278," Volume 1, CM GNCS Operations. Mass. Instit. of Tech./Inst. Lab report R-547, October 1966.

9. Peters, W. H.: "Coupling Between Elastic Airframe Resonances and Rocket Engine Gimbal Actuation." NASA MSC memorandum EG23, September 23, 1965.

10. Knobbe, E. J.: "Block II Autopilot Analysis LM-ON." North American Rockwell Internal Letter FS/GCA/65/115, July 14, 1965.

11. Peters, W. H.: "In-flight Stroking Tests for Apollo." NASA MSC memorandum EG23-132-65-906, October 21, 1966.

12. "Action Items and Agreements of Powered Flight Design Review Meeting." Mass. Instit. of Tech./Inst. Lab report Minutes of Meeting held December 13, 1967.

13. Peters, W. H.: "Minutes of TVC DAP Design Review Meeting." NASA MSC memorandum EG23-67-119, July 7, 1967.

14. Peters, W. H.: "Official Apollo Bending Mode Shapes and Tolerances." NASA MSC memorandum EG23-67-168, September 11, 1967.

15. Peters, W. H.: "Dynamic Testing of Apollo Spacecraft in the Flight Environment (Stroking Test)." NASA MSC Internal Note MSC-EG-68-19, December 31, 1968.

16. Marcantel, Bernard: "Detailed Post Flight Analysis of Data from Apollo 9 In-flight Dynamics Test (Stroking Test)." NASA MSC Internal Note MSC-EG-69-34, September 4, 1969.

BLANK PAGE

ORBITING ASTRONOMICAL OBSERVATORY (OAO-A2) STABILIZATION AND CONTROL SUBSYSTEM

By Thomas E. Huber
Goddard Space Flight Center

SUMMARY

This paper describes the OAO-A2 control system and the flight results to date. The OAO (Figure 1) has been successfully operating for approximately 11 months and has made over 2500 stellar observations. The OAO control system is a complex system which takes the 4500-lb vehicle through initial stabilization to star tracker control where a pointing accuracy of less than 1 arc minute is achieved. To obtain this pointing accuracy the system uses six orthogonally mounted gimbaled star trackers. The trackers are two gimbaled devices having a 1-deg field of view with a ±43-deg gimbal excursion. The tracker can recognize and track 2.0 magnitude stars or brighter. The flight results indicate spacecraft pointing accuracies of less than 1 arc minute after tracker calibration with spacecraft jitter less than 3 arc seconds.

INTRODUCTION

Initial Stabilization and Roll Search

The initial stabilization of the spacecraft is accomplished by coarse and fine sun sensors driving a jet system and fine inertia wheels. The roll axis is aligned to the sun line within 0.25 deg. Rate gyros are used for damping during sun acquisition. To acquire the stars, a roll search program is initiated. The program consists of inserting a fixed bias into the roll rate gyro channel which causes the vehicle to rotate about the sun line at a fixed rate (0.25 deg/sec). The trackers are prepositioned such that at some roll attitude about the sun line each assigned tracker will have a guide star enter its field of view.

When the trackers simultaneously have star presence signals, they switch from command to track mode and track the star. When the tracking errors are reduced to less than 2 arc minutes, the roll rate bias is removed. When a roll rate null is achieved, the system is switched to star tracker control.

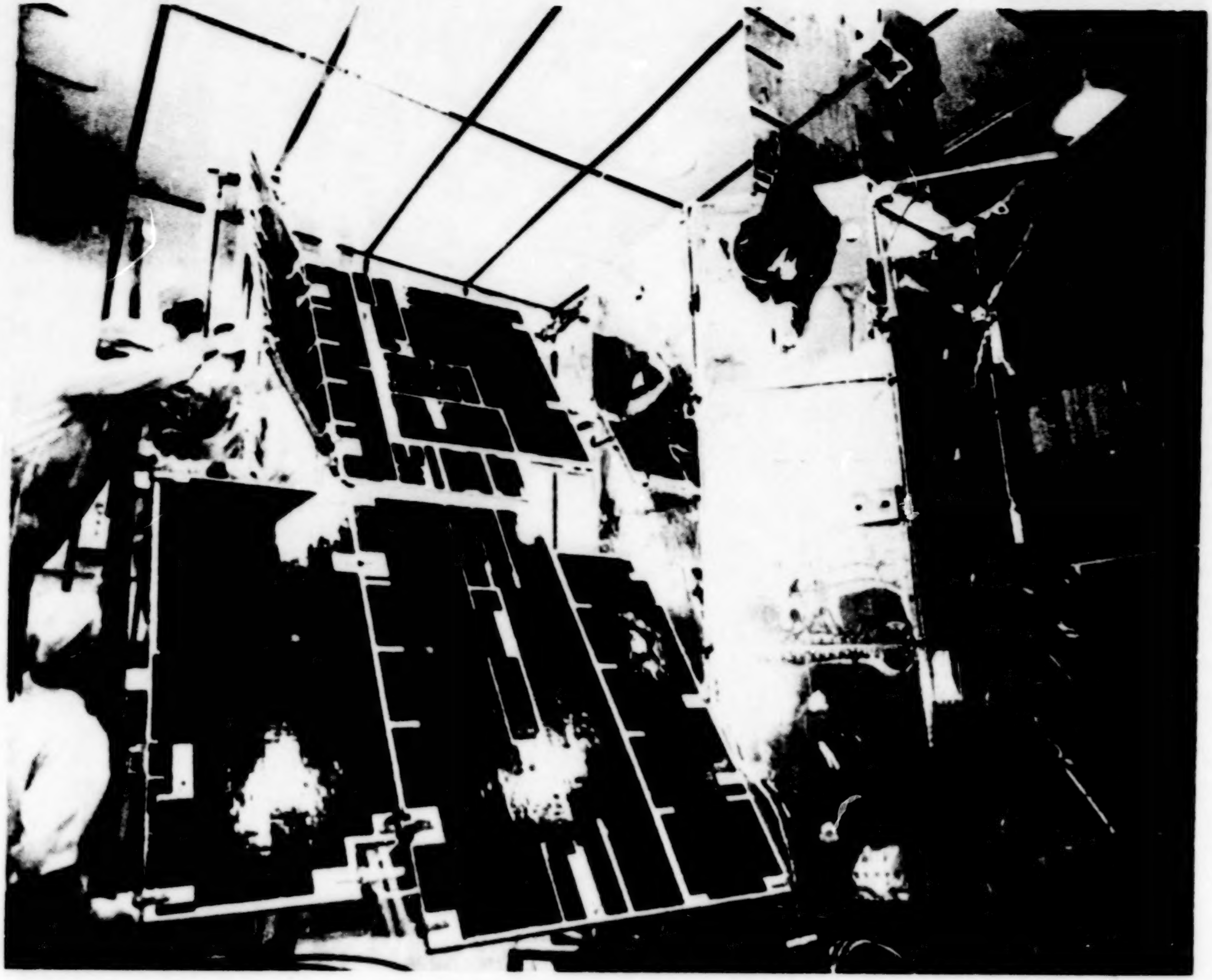


Figure 1.- OAO

Sun Bathing

To automatically acquire a power positive condition, a sun bathing mode of operation was incorporated in OAO-A2. A set of sun sensors mounted on the solar paddle axes drive the spacecraft through a jet system to an orientation where the solar paddles are normal to the sun. Rate integrating gyros are used for damping during sun acquisition. The gyros automatically switch from rate mode to attitude mode when the spacecraft enters darkness. The system then maintains attitude reference by driving the jet system via the gyro attitude errors. The flight data show that the drift rate of the gyro system is less than 0.20 deg/hr, which is better than the spec value of 0.3 deg/hr. The gyro errors can be switched to drive the fine inertia wheels through the fine wheel and jet controllers. This holding mode of operation saves gas and provides a stable reference to perform star search.

Slewing

The spacecraft is slewed from one experiment star to another by a coarse wheel system. A slew command is loaded into the spacecraft which energizes the appropriate coarse momentum wheel (pitch, roll, or yaw). A pulse is generated from the wheel per revolution which counts down the slew command counter. When the counter reaches zero, a dc brake is applied to the wheel and terminates the slew. Two or more trackers are required to be tracking stars during the slew maneuver. The tracker command angles are updated during the slew and the tracker system will reduce the spacecraft error to one arc minute after a settling period when the slew is completed.

Momentum Unloading

The momentum unloading of the fine wheels can be accomplished by two methods:

1. Gas unloading
2. Magnetic unloading

Gas unloading utilizes a low thrust jet system which operates automatically when the fine wheel speeds reach 75 percent of no-load wheel speed. The jet system can also be actuated by transmitting 5 percent or 40 percent unload commands. If the wheel speeds exceed these values, the jets will fire and unload the wheels to approximately 8 percent.

The Magnetic Unloading System (MUS) continuously dumps the momentum of the fine wheels. The system consists of three magnetometers, a magnetic signal processor, and three torquer bars. The signal processor sums the components of earth's field sensed by the magnetometers and the errors generated by the sensors and applies a proportional current to the torquer bars. This current generates a field which reacts with the earth's field producing a torque that effectively unloads the wheels.

The system is capable of generating 300 dyn-cm of torque. When operated in orbit, the MUS has kept the wheel's speeds below 20 percent throughout the flight.

Star Search

The OAO has lost stabilization 30 times to date. The stars have been re-acquired 29 times through a star search program. The remaining acquisition was accomplished by a normal roll search. The star search routine is performed during the gyro hold-on-wheels mode. Star search consists of determining the spacecraft attitude by ground processing Adcole solar aspect data and magnetometer data and deriving a weighted least squares fit. After ground processing, tracker gimbal command angles are transmitted to the spacecraft, which move the tracker gimbals through a star search pattern based on the computed attitude. When the guide star is acquired, the tracker switches automatically from command to track mode and tracks the star. When the assigned trackers have acquired their guide star, the system is switched to tracker control.

Symbols and Abbreviations

CSS - Coarse Sun Sensors
FSS - Fine Sun Sensors
DE - Disable Eye
JRT-45 - Rate Gyros
SSP - Sensor Signal Processors
FW & JC - Fine Wheel and Jet Controller
F/W - Fine Wheel
HTJ - High Thrust Jets
CW - Clock-wise
CCW - Counter Clock-wise
RAPS - Rate and Position Sensor
PPDS - Primary Processor and Data Storage
TRK - Tracker
PMT - Photo-Multiplier Tube
BST - Bore Sight Tracker
DLU - Digital Logic Unit

TABLE OF CONTENTS

1. SPACE VEHICLE CONTROL APPLICATIONS

Design Development of the Apollo Lunar Module
Kenneth J. Cox
Manned Spacecraft Center 1 1/A10

Design Development of the Apollo Command and Service Module Thrust Vector Attitude Control System
William H. Peters
Manned Spacecraft Center 63 1/E3

Orbiting Astronomical Observatory (OAO-A2) Stabilization and Control Subsystem
Thomas E. Huber
Goddard Space Flight Center 83 1/G10

Applications of the Hybrid Coordinate Method to the TOPS Autopilot
G. E. Fleischer
Jet Propulsion Laboratory 127 2/D7

Steering Laws for Double-Gimbal Control-Moment Gyros
Stephen W. Winder
Marshall Space Flight Center 157 2/F13

2. AIRCRAFT GUIDANCE, CONTROL AND HANDLING QUALITIES

Optimum Horizon Guidance Techniques for Aircraft
Heinz Erzberger and Homer Q. Lee
Ames Research Center 175 3/A8

Inertially-Aided, All-Weather Landing Systems for C/TOL.
Robert J. Pawlak
Electronics Research Center 185 3/B4

An Approach to the Determination of Aircraft Handling Qualities Using Pilot Transfer Functions
James J. Adams and Howard G. Hatch, Jr.
Langley Research Center 199 3/C4

A Frequency Domain Approach to Handling Qualities Design
William A. Wolovich
Electronics Research Center 217 3/D8

3. SYSTEM SIMULATION AND IDENTIFICATION

Optimal Regulator or Conventional? Setup Techniques for a Model Following Simulator Control System
Dwain A. Deets
Flight Research Center 237 3/F1

TABLE OF CONTENTS (CONT'D)

3. SYSTEM SIMULATION AND IDENTIFICATION (CONT'D)

Quasi-Optimal Control of a Moving-Base Simulator
Elwood C. Stewart
Ames Research Center 253 3/G3

Some Modern Control Techniques for Human Operator Modeling and Identification
Richard S. Shirley
Electronics Research Center 261 3/G11

4. ENGINE CONTROL AND ADVANCED PROPULSION TECHNIQUES

Normal Shock and Restart Controls for a Supersonic Airbreathing Propulsion System
George H. Neiner, Gary L. Cole, and
Francis J. Paulovich
Lewis Research Center 299 4/D2

Optimal Control of a Supersonic Inlet to Minimize Frequency of Inlet Upstart
Bruce Lehtinen, John R. Zeller, and
Lucille C. Geyser
Lewis Research Center 323 4/F1

Control System Studies for Thermionic Reactors
Richard J. Hermsen,
California State Polytechnic College
Henrik G. Gronroos
JPL, California Institute of Technology 337 4/G1

Application of Fluidics to New Control Components
Miles O. Dustin, Vernon D. Gebben, and
Robert E. Wallhagen
Lewis Research Center 365 5/B7

Minimum Energy Control of a Class of Electrically Driven Vehicles
Yilmaz E. Sahinkaya
JPL, California Institute of Technology 387 5/D6

5. ADVANCED CONTROL TECHNIQUES

A New Formulation for the Epsilon Method Applied to the Minimum Time-to-Climb Problem
Lawrence W. Taylor, Jr., and Harriet J. Smith
Flight Research Center 423 5/G1

TABLE OF CONTENTS (CONT'D)

5. ADVANCED CONTROL TECHNIQUES (CONT'D)

Determination of Optimal Control Using Imbedding of the Terminal Conditions
Raymond C. Montgomery
Langley Research Center **435** 5/G13

Stability of Constant Gain Systems with Vector Feedback
George L. von Pragenau
Marshall Space Flight Center **443** 6/A10

Synthesis of Finite Settling Time Discrete Systems
James A. Gatlin
Goddard Space Flight Center **473** 6/C12

Design and Performance of Heart Assist or Artificial Heart Control Systems
John A. Webb, Jr., and Vernon D. Gebben
Lewis Research Center **495** 6/E1

STSP - Star Tracker Signal Processor
GST - Gimbal Star Tracker
TPC - Tracker Power Controller
BSTLU - Bore Sight Tracker Logic Unit
MUS - Magnetic Unloading System
WEP - Wisconsin Experiment Package
SAO - Smithsonian Astrophysical Observatory
GEP - Goddard Experiment Package
PEP - Princeton Experiment Package
SCPS - Support Computer Program System

DISCUSSION

On December 7, 1968, the first successful Orbiting Astronomical Observatory (OAO-A2) ever launched was placed in a 500-mile circular orbit above the earth. It carried two experiments: the Wisconsin Experiment Package (WEP) and the Smithsonian Astrophysical Observatory (SAO). The WEP has seven separate photometric systems which study the stars in the 1000 to 3000-Å region. WEP is making approximately 13 observations/day and has made over 2500 observations to date. The SAO measures the brightness of young stars in four spectral bands between 1000 and 3000 Å. The instruments used are called the Celescope (celestial telescope).

The Celescope measures with four large-aperture television cameras using broad-band television photometers. A special type of television tube, called the "Uvicon", sensitive only to ultraviolet light, was developed for Celescope. The SAO has taken approximately 5000 pictures and 1700 observations to date.

To perform these observations, a complex stabilization and control was developed. The mission requirements of the Stabilization and Control Subsystem are:

1. To acquire external attitude references.
2. To stabilize the Observatory to these references for the purpose of providing a pointing reference for the experimental package.
3. To maneuver the Observatory to any space attitude by open loop control and settle under control of star trackers.
4. To maintain the Observatory attitude under control of inertial devices when required.
5. To re-orient the Observatory under control of sun sensors for purposes of efficient solar power collection when required.

The external references chosen for the control system are the sun and 34 specified stars.

The Stabilization and Control System has basically seven modes of operation:

1. Initial stabilization
2. Sun bathing
3. Roll search
4. Star search
5. Coarse star pointing
6. Re-orientation
7. Fine pointing (flights 3 and 4 only).

Each mode will be described in detail and the OAO-A2 flight results will be discussed. Figure 2 shows the first three modes of operation.

STABILIZATION AND CONTROL MODE DESCRIPTION

Initial Stabilization

Rate stabilization and coarse solar acquisition.- Initial stabilization of the OAO spacecraft is to align the negative roll axis ($-X_C$) (see Figure 3) of the spacecraft to the sun line. This orientation is called beta equal zero. It is accomplished by using coarse and fine solar sensors and rate gyros.

The coarse sun sensors are silicon solar cells mounted on the vehicle to cover a 4-π sr field of view. These sensors exhibit a response to sun light that is approximately cosine response. There are a total of eight coarse eyes on the vehicle. Four eyes are used to sense errors about the spacecraft pitch (Y_C) control axis, and four are used to sense errors about the yaw (Z_C) axis. No control is provided about the roll (X_C) axis by the solar sensors. Two of the four eyes for each axis are located at the sun pointing (aft) end of the OAO. These cells are set at an angle of 50 deg to the sun line ($-X_C$) axis pointing in opposite directions. The polarity of one cell is reversed with respect to the other cell and the outputs of the two cells are hard wired together and output currents add algebraically. A second pair of coarse eyes are positioned at the front end of the OAO. These eyes are so connected that they reinforce the output of the rear mounted eye on that side of the spacecraft. The angle of each of these front eyes to the sun line is 30 deg to minimized errors due to the earth's albedo.

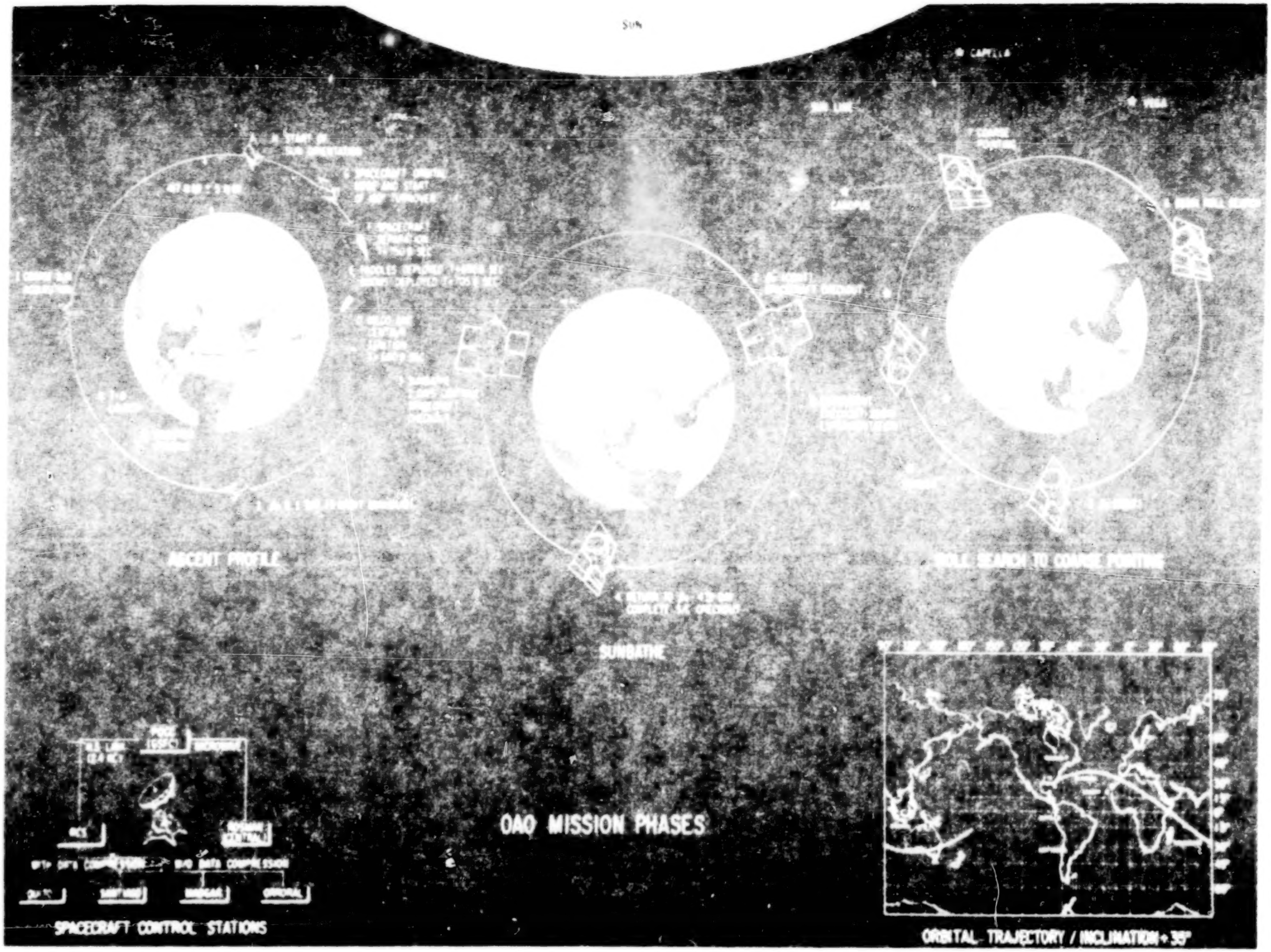

SUN
CAPELLA
VEGA
SUN LINE
CANOPUS
ASCENT PROFILE
SUNBATHE
ROLL SEARCH TO COARSE POINTING
OAO MISSION PHASES
POCC
(GSFC)
SPACECRAFT CONTROL STATIONS
ORBITAL TRAJECTORY / INCLINATION = 35°

Figure 2.- OAO mission phases

54

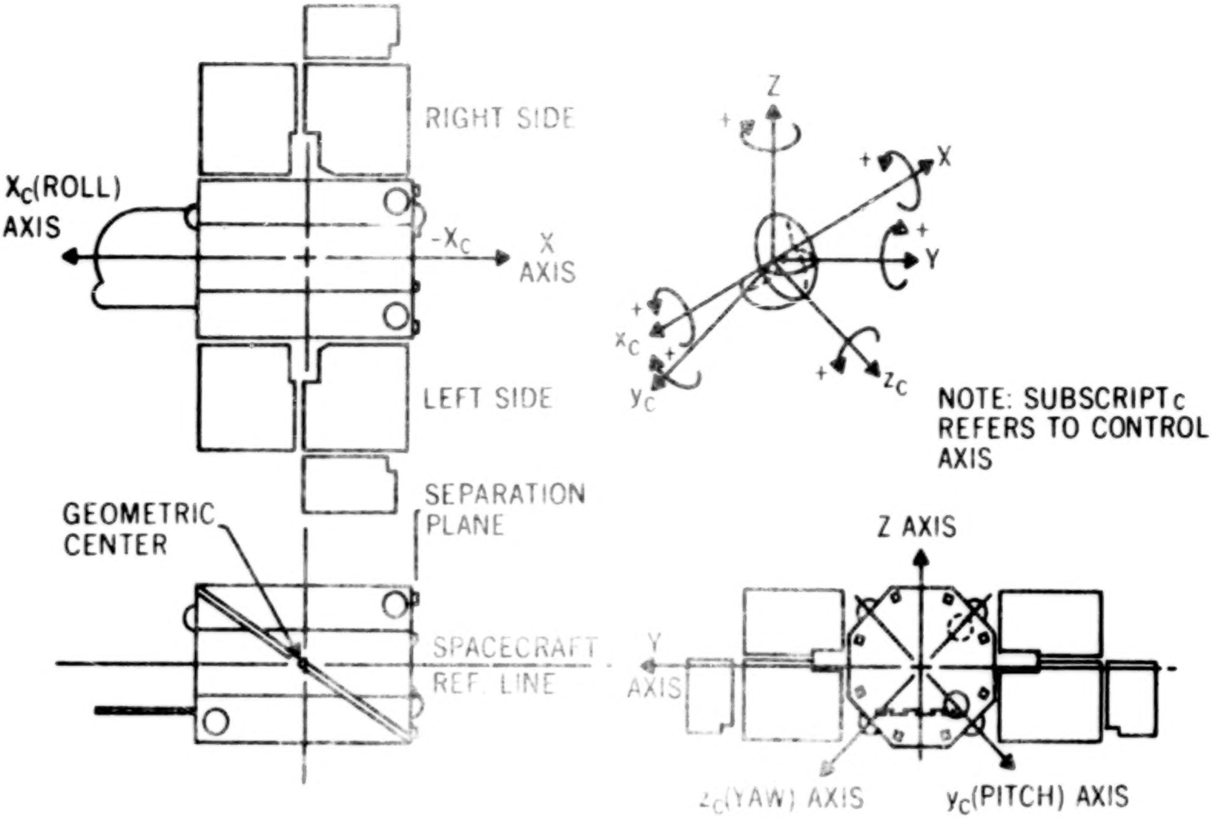

RIGHT SIDE
X_c(ROLL) AXIS
$-X_C$
X AXIS
LEFT SIDE
Z
X
Y
X_C
Y_C
Z_C
NOTE: SUBSCRIPT c REFERS TO CONTROL AXIS
GEOMETRIC CENTER
SEPARATION PLANE
SPACECRAFT REF. LINE
Y AXIS
Z AXIS
Z_C(YAW) AXIS
Y_C(PITCH) AXIS

Figure 3.- Coordinate reference system

The signal flow (Figure 4) in the pitch and yaw channels is from the rate gyros (JRT-45) and Coarse Sun Sensors (CSS) to the Sensor Signal Processors (SSP's) to the Fine Wheel and Jet Controllers (FW&JC's) and then to the appropriate high thrust jets. All observatory torquing during this mode is accomplished using high thrust gas jets. The system is capable of operating from any spatial orientation with initial observatory rates of up to 1.0 deg/sec.

The control jets reduce the angular rates about the pitch and yaw axis to within 0.05 deg/sec within 10 min after separating with initial rates of 1 deg/sec. In this mode, the control system is capable of pointing the negative roll axis of the observatory to the sun line within 13 deg. This value includes worst case errors due to the earth albedo.

The design control law for the coarse sun mode is:

$$\theta + 7\dot{\theta} = \pm 2^\circ$$

Fine solar acquisition. After completion of rate stabilization and coarse solar acquisition, the control system further reduces the pointing error to the sun. The Disable Eye (DE),

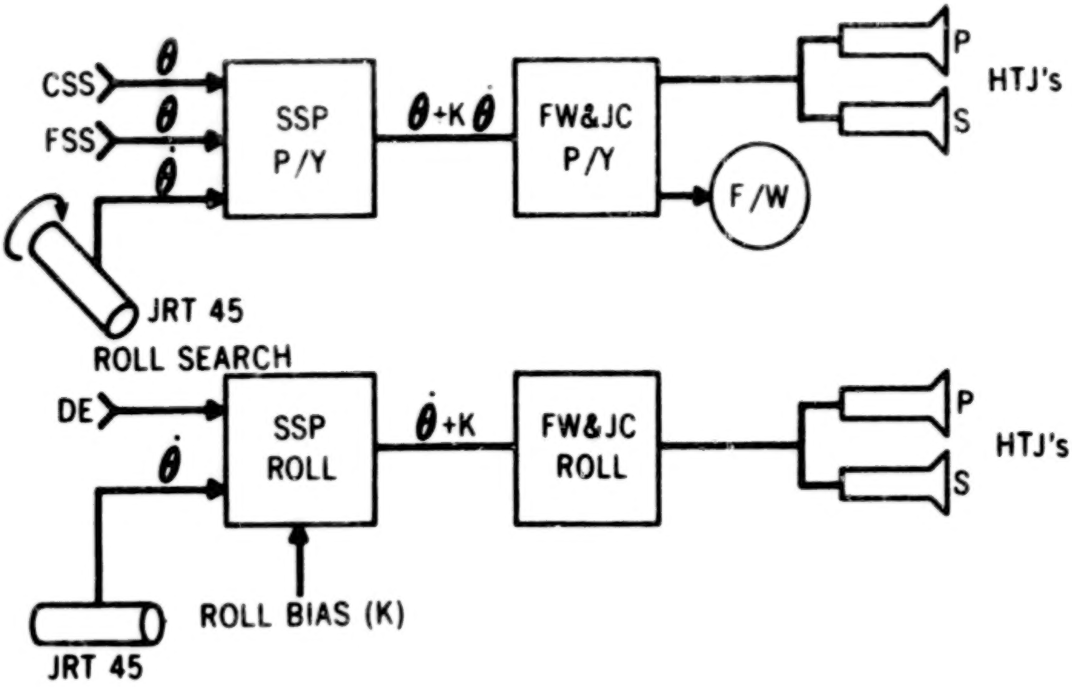


Figure 4.- Initial stabilization block diagram

with a conical field of view, senses when the sun is within 13 deg of the negative roll axis and produces a bi-level output. Null detectors in the pitch and yaw SSP's sense when the angular rates are 0.06 deg/sec or less. These three bits of information are sent to a programmer as logic signals where they are processed through an "And" gate whose output is then fed to the pitch and yaw FWJC's. In the FWJC's this signal switches the pneumatics channel inputs from the coarse sun sensor error signals to the fine sun sensor error signals, enables the pitch and yaw fine wheels and enables the drivers of the low thrust jets whose function is to unload the fine wheels. The signal flow (Figure 4) in the pitch and yaw channels is from the rate gyros and fine sun sensors to the sensor signal processors to the fine wheel and jet controllers and then to the appropriate high thrust jets. In addition, the fine sun sensor signals drive the Fine Wheels (FW). Signal flow in the roll channel is the same as described for rate stabilization in the coarse sun mode. The Observatory's negative roll axis is pointed at the sun to within 0.25 deg. This tolerance includes the sum of all errors introduced by the equipment and the Observatory thermal and alignment errors.

The design control law for this fine sun mode is:

$$\theta + 1.16\,\dot{\theta} = 0.333^{\circ}$$

Initial stabilization flight performance.- Although originally conceived as a preliminary mode from which roll search could be performed and control transferred to gimbal trackers, initial stabilization of A2 was to provide a stable orientation at beta zero for thermal cool-down of the spacecraft during the first four orbits.

Coarse solar stabilization.- Coarse solar stabilization was initiated approximately 12 min after launch. All systems functioned properly with initial separation rates and attitude as follows:

Roll rate = -.01 deg/sec
Pitch rate = -.09 deg/sec
Yaw rate = +.16 deg/sec

Pitch attitude error = +76 deg
Yaw attitude error = +104 deg

Computer studies had estimated coarse stabilization settling times varying from 6 to 9 min,depending on the initial pitch and yaw attitude errors, and gas consumption on the order of 1 lb. Actual coarse stabilization settling required 5.9 min and only .514 lb of gas was consumed. The unusually low expenditure was due primarily to extremely low separation rates. Tip off rates on the order of 1 deg/sec were anticipated but did not occur.

Pitch and yaw coarse stabilization phase plane trajectories are plotted in Figures 5 and 6. These figures show that the actual high thrust jet control laws never varied by more than 15 percent from nominal and were well within acceptable dynamic requirements.

Flight fine solar stabilization.- Approximately 5.9 min after initial separation, simultaneous pitch and yaw rate nulls and disable eye signal were obtained and the spacecraft automatically transfered to the fine solar stabilization phase of initial stabilization.

The conditions of transfer were:

Roll rate = .01 deg/sec
Pitch rate = -.004 deg/sec
Yaw rate = .06 deg/sec

Pitch attitude error = +1.3 deg
Yaw attitude error = +2.2 deg

All systems functioned properly during this phase, 206 sec and .233 lb of gas were required to successfully complete the stable alignment of the negative roll axis to the sun line. These parameters compare favorably to the estimated computer settling time of 4 to 6 min and gas consumption of approximately 0.5 lb.

Pitch and yaw phase plane trajectories for fine solar stabilization are shown in Figures 7 and 8. Variations from the nominal design control laws are negligible.

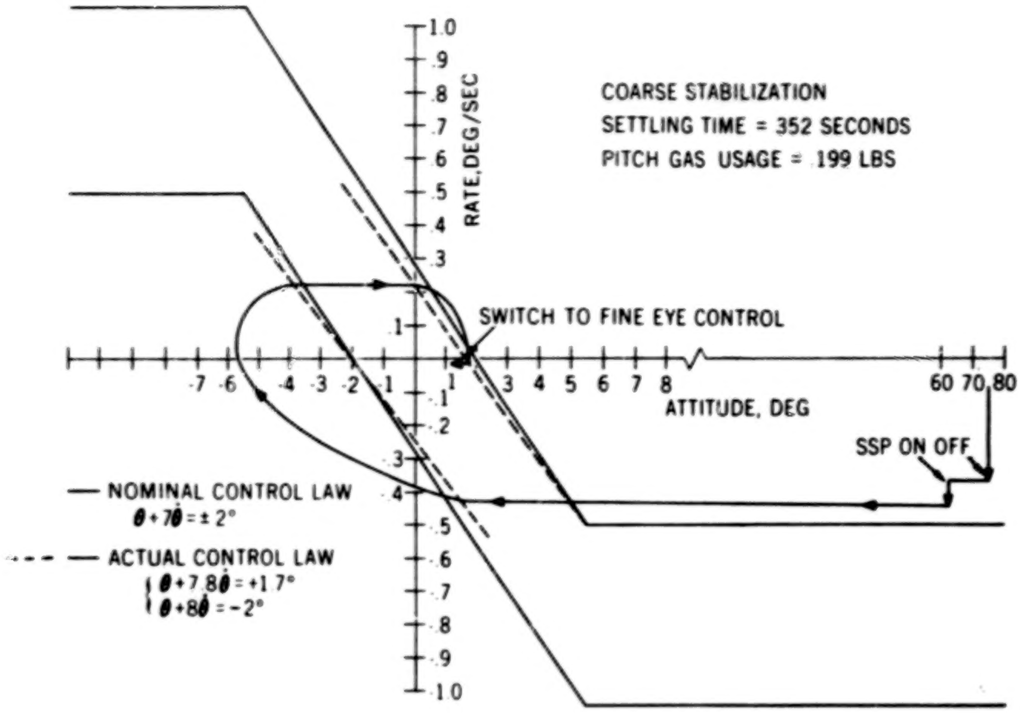


Figure 5.- Pitch rate vs attitude phase plane, coarse stabilization

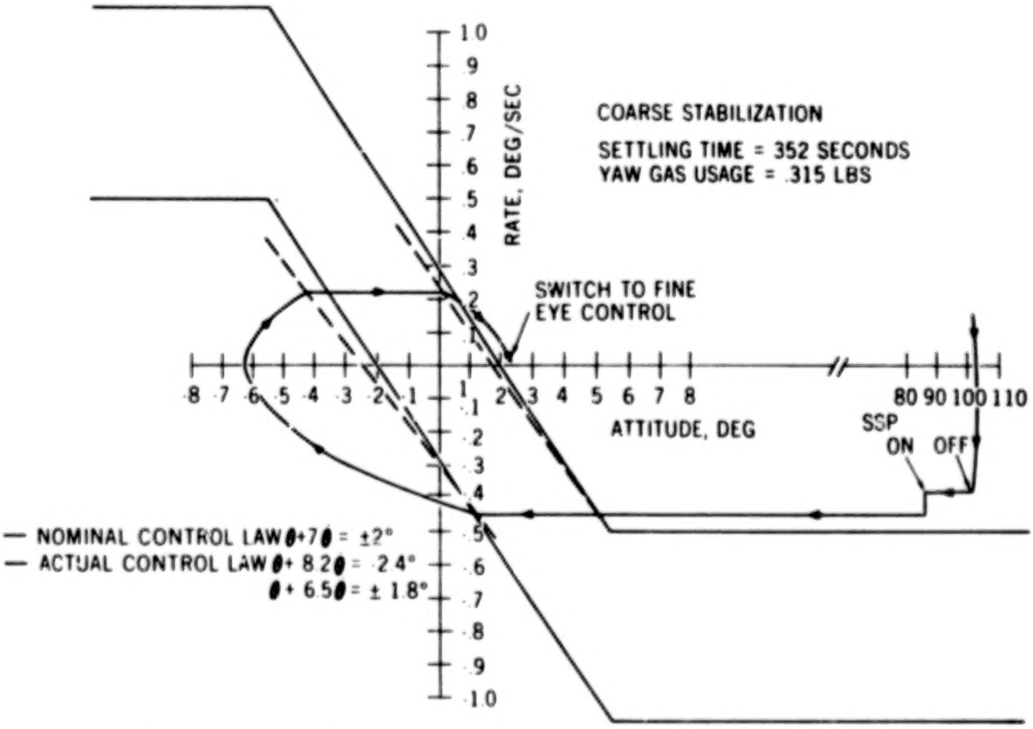


Figure 6.- Yaw rate vs attitude phase plane, coarse stabilization

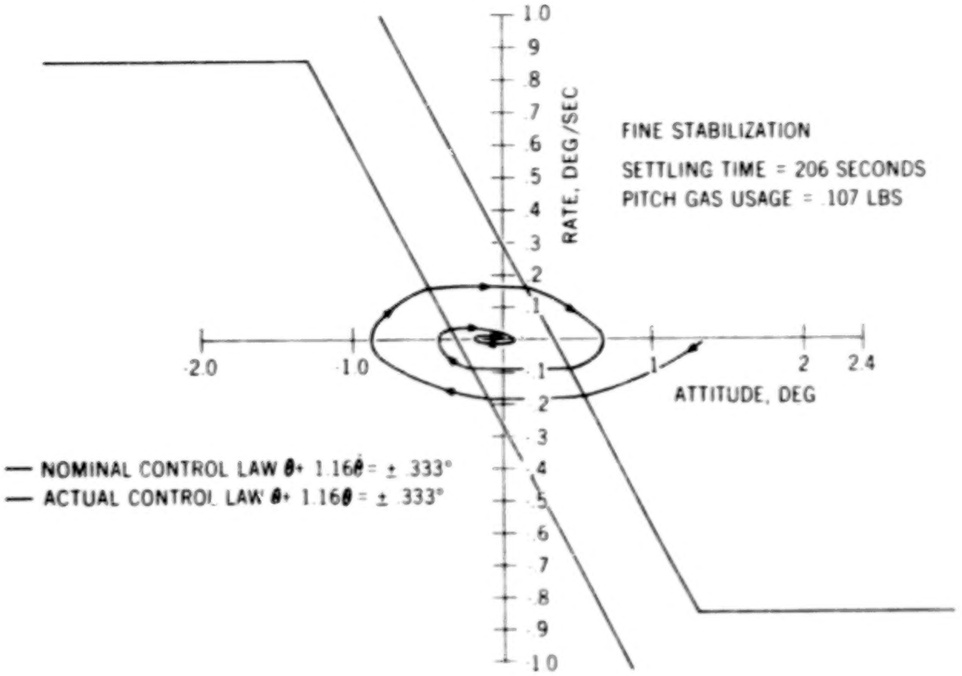


Figure 7.- Pitch rate vs attitude phase plane, fine stabilization

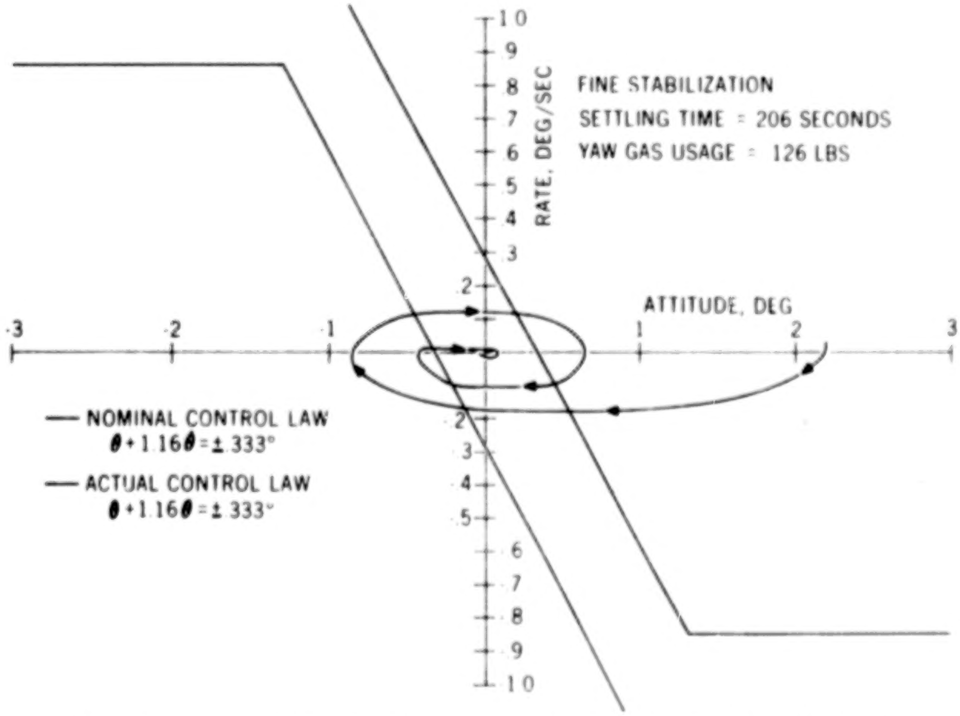


Figure 8.- Yaw rate vs attitude phase plane, fine stabilization

Time history plots of pitch and yaw rate, attitude and wheel speeds, demonstrating fine wheel capture at the termination of fine solar stabilization are shown in Figures 9 and 10.

Sun Bathing

The sun-bathing mode of operation is to orient the OAO spacecraft so that the solar paddles (B side) are pointed normal to the sun (Figure 11). The initial checkout of the A2 vehicle was conducted in this mode of operation. The sun-bathing sequence is stored in the spacecraft memory and is executed upon loss of star tracker control or if a critical under voltage condition is generated.

RAPS subsystem.- To implement sun bathing, a Rate and Position Sensor (RAPS) system was incorporated on A2. The system consisted of four separate components (Figure 12):

1. Gyro unit
2. Gyro electronics
3. RAPS sun sensors
4. RAPS controller.

Gyro unit: The gyro unit consists of three Kearfott Alpha Rate Integrating gyros, gyro heaters, and thermistors. Each gyro is capable of producing either attitude or rate information depending upon whether or not the loop is closed between the gyro pickoff and the torquer through the gyro electronics. When the loop is closed, rate mode is established. The attitude mode operation is achieved with the switch open. The function of the RAPS is to provide signal proportional to the angular velocities or attitudes of the OAO about its principal control axis.

Gyro electronics: The electronics consist of logic modules and torque motor amplifiers. The unit also contains:

1. Temperature regulating circuitry to maintain gyros at $140^{o}F \pm 2^{o}F$
2. Telemetry
3. An 800-cps reference generator for the gyro pickoff.

RAPS sun sensors: The RAPS system has a separate array of sun sensors. The RAPS sun sensors provide attitude information about the solar paddles axes. There are eight sun sensors:

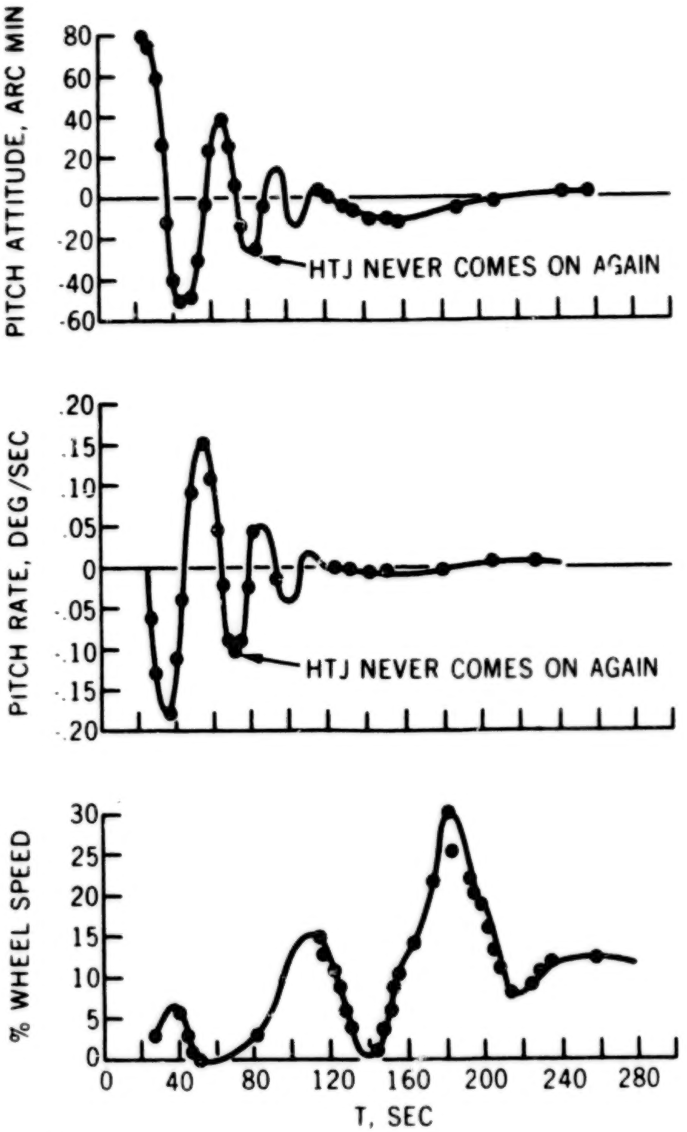


Figure 9.- Pitch fine wheel capture during initial stabilization

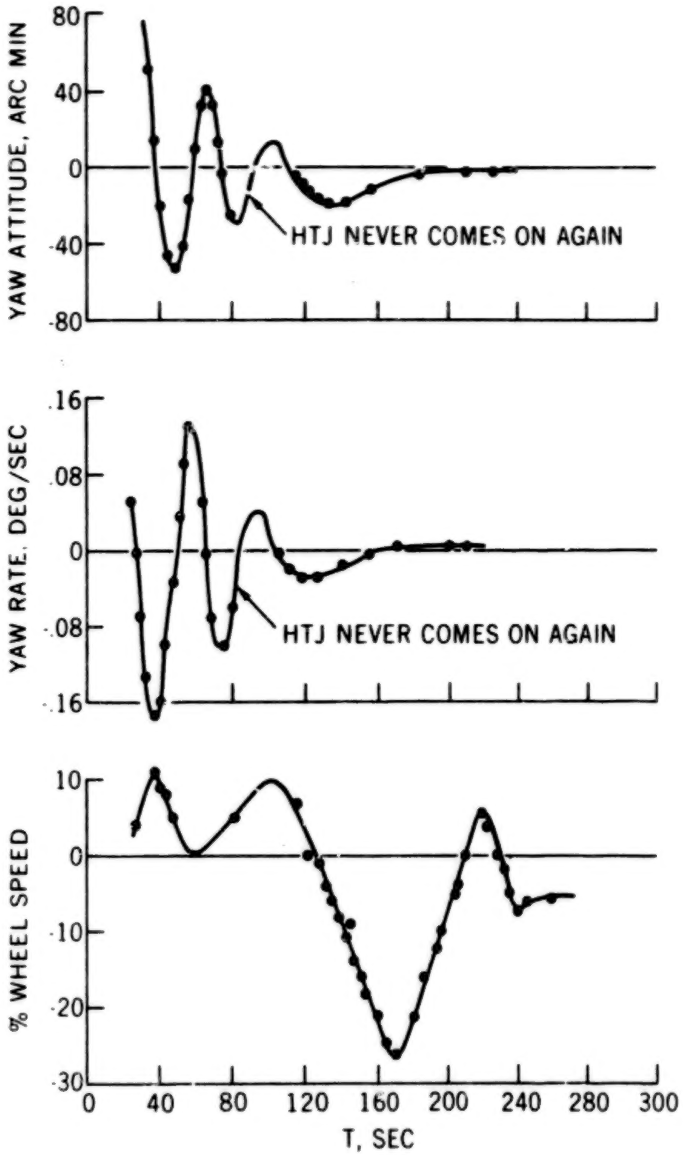


Figure 10.- Yaw fine wheel capture during initial stabilization

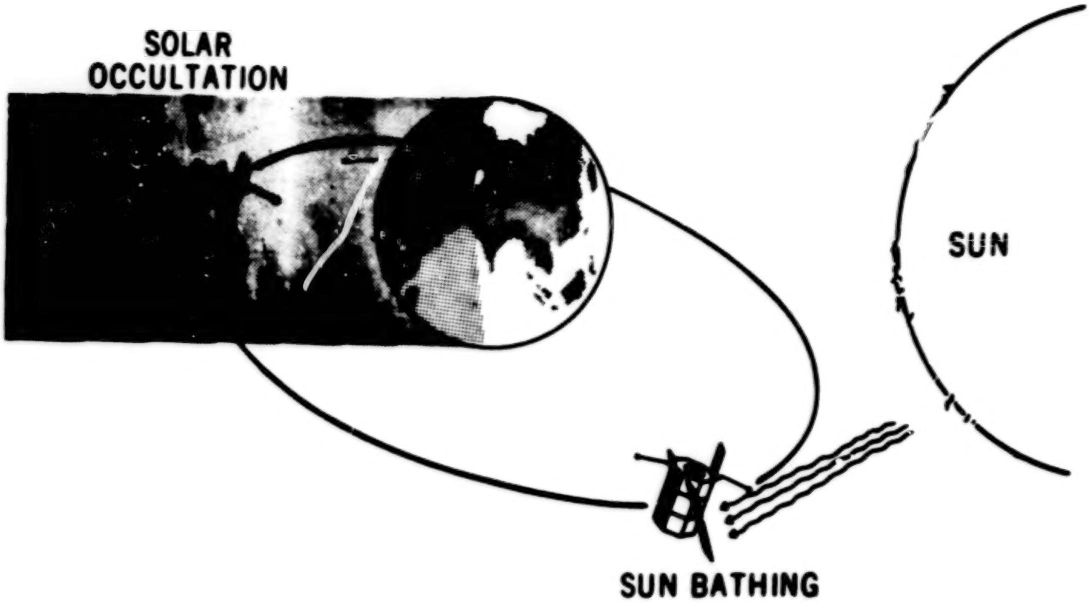

SOLAR OCCULTATION
SUN
SUN BATHING

Figure 11.- OAO-A2 rate and position sensing system

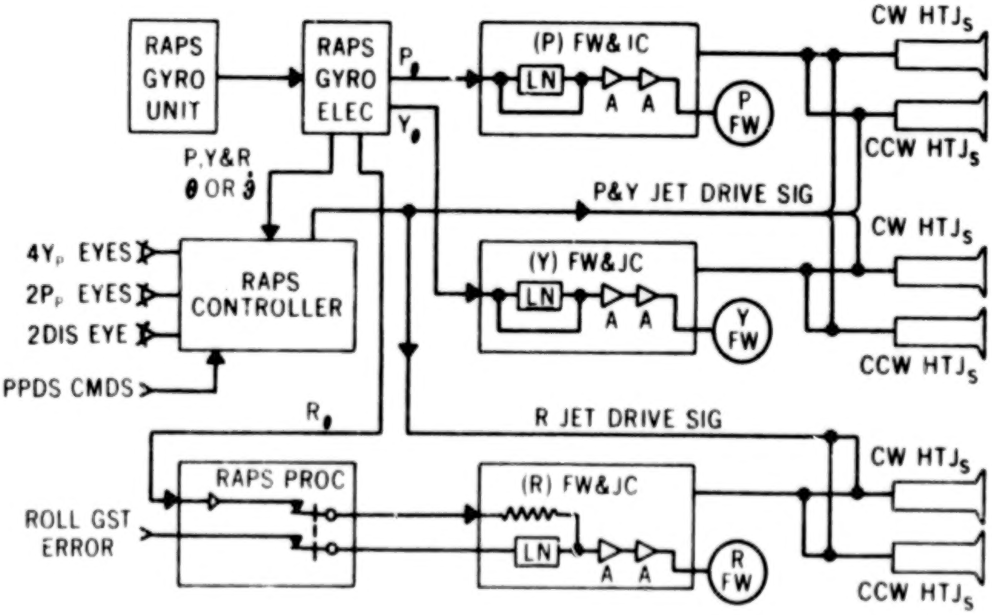

RAPS GYRO UNIT
RAPS GYRO ELEC
P_θ
Y_θ
(P) FW&JC
LN
A A
P FW
CW HTJs
CCW HTJs
P,Y&R
θ OR θ̇
P&Y JET DRIVE SIG
CW HTJs
4Y_P EYES
2P_P EYES
2DIS EYE
RAPS CONTROLLER
(Y) FW&JC
LN
A A
Y FW
CCW HTJs
PPDS CMDS
R_θ
R JET DRIVE SIG
CW HTJs
RAPS PROC
(R) FW&JC
ROLL GST ERROR
LN
A A
R FW
CCW HTJs

Figure 12.- RAPS subsystem

Yaw Set - Four sensors, arranged in a null pair and anti-null pair.

Pitch Set - Two sensors, arranged in a null pair.

Disable Set - Two sensors, arranged diametrically opposite each other.

The pitch and yaw sets sense the direction of the sun with respect to the normal to the solar paddles. The disable set senses the presence of the sun.

RAPS controller: The RAPS controller accepts the gyro and RAPS sun sensor signals, operates on them, and generates drive signals to operate the secondary high thrust jets. Since the jets are mounted on the control axes (X_C, Y_C, Z_C) and the sensors are mounted on the solar paddle axes (X_p, Y_p, Z_p), a transformation matrix must take place in the RAPS controller.

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} -.577 & .833 & 0 \\ -.588 & -.394 & -.707 \\ -.588 & -.394 & .707 \end{bmatrix} \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix}$$

Therefore, the control axes in terms of paddle axes are:

$$X_c = 0.833\, Y_p$$

$$Y_c = -0.394\, Y_p \quad -0.707\, Z_p$$

$$Z_c = -0.394\, Y_p \quad +0.707\, Z_p$$

Sun bathing operation: When sun bathing mode occurs, the RAPS controller receives attitude errors from the RAPS sun sensors and rate errors from the RAPS gyros and combines them to drive the appropriate high thrust jet to orient the spacecraft. The design control law for this mode is:

$$\theta + 20\,\dot{\theta} = \pm 1^{\circ}$$

A stabilized sun hold attitude on jets is obtained during the sunlit portion of the orbit with the following error limits:

Less than 5 deg about the paddle pitch axis.

Less than 7 deg about the paddle yaw axis.

During the dark portion of an orbit, the RAPS controller automatically switches the gyros to their attitude mode when signals from the RAPS sun sensors indicate the sun is not present. The design control law using gyro attitude information driving appropriate jets is:

$$\theta + 9\dot{\theta} = \pm 0.5^{\circ}$$

The RAPS gyro attitude information can be fed to the fine wheels to obtain a gyro hold-on-wheel mode. This mode is commanded to save gas once the system has settled in the sun-bathing mode. It is also used as a hold mode during guide star occultations. The random drift rate of gyros was spec at 0.3 deg/hr. The hold mode of operation was designed to operate for approximately ten orbits before updating is required by switching back to RAPS sun sensor control.

Sun bathing and RAPS flight performance.- Figures 13, 14, and 15 show that the performance of the RAPS system was nominal when sun bathing was commanded during the fourth orbit.

As shown by Figure 16, the RAPS drive was approximately a factor of two or better than specification on all axes. The stability of the gyros approached 0.03 deg/hr. On the third OAO spacecraft, a precision inertial reference unit will be flown built by MIT. The system will be implemented to have drift updating capability. The random drift can be compensated by ground command and an inertial system with a drift rate of 0.003 deg/hr will result.

The RAPS contribution to the success of OAO-A2 was truly significant. It provided:

1. "Parking" mode where the spacecraft was power positive.
2. A stable reference where a star search routine could be performed.
3. A satellite holding function to point to areas of the celestial sphere where adequate guide stars for three axis control were not available.

RAPS failure. — On August 21, 1969 (nine months after launch) during Orbit 3701, the RAPS pitch gyro failed. A drift test was being conducted when it was noticed that the pitch axis drift exceeded 1 deg/hr. Roll and yaw axis drift were normal. The resultant trouble-shooting indicated that the gyro would not correctly respond to spacecraft motion in either rate or attitude mode. The (H) of the gyro had greatly reduced and

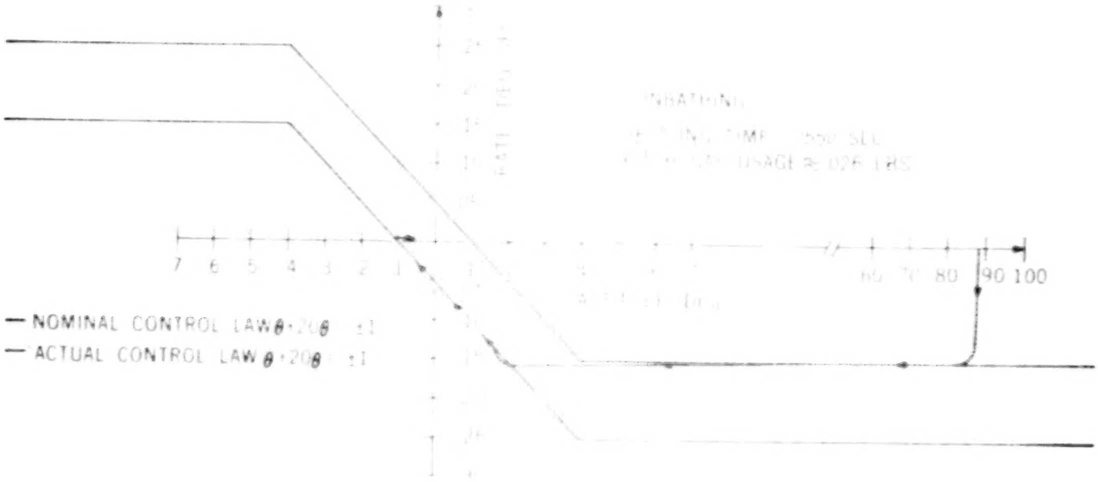


Figure 13.- Pitch rate vs attitude phase plane sun bathing

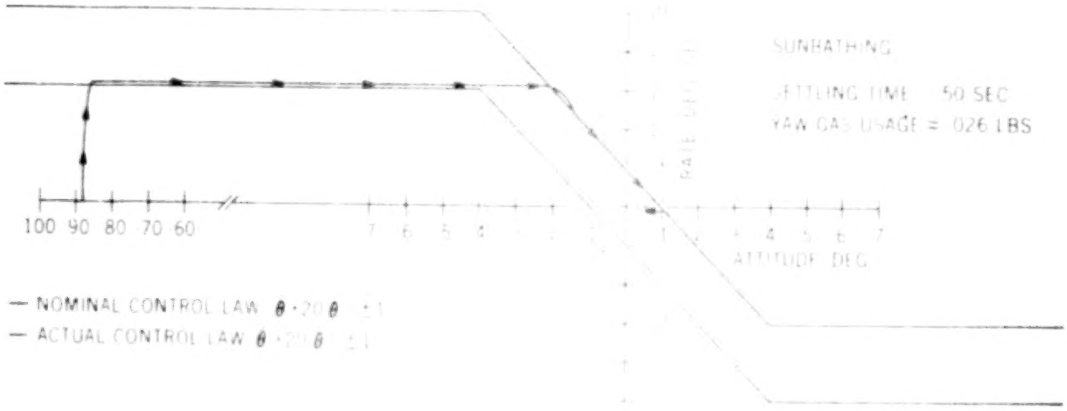


Figure 14.- Yaw rate vs attitude phase plane sun bathing

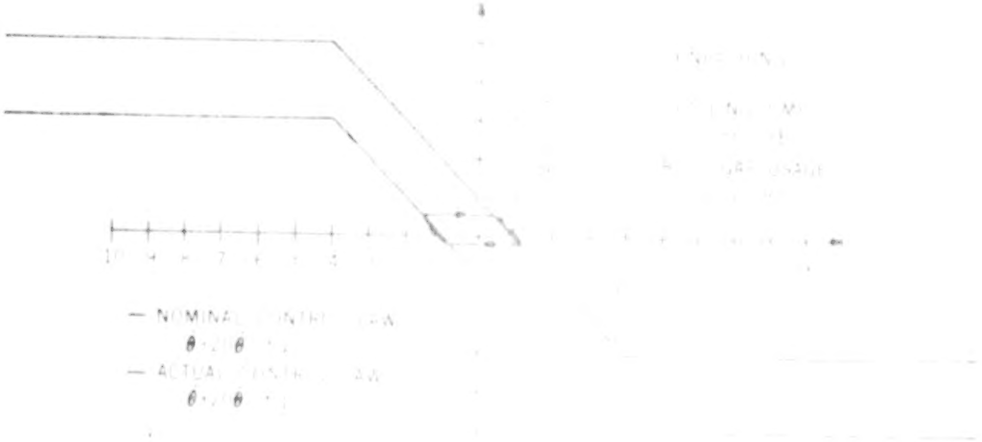


Figure 15.- Roll rate vs attitude phase plane sun bathing

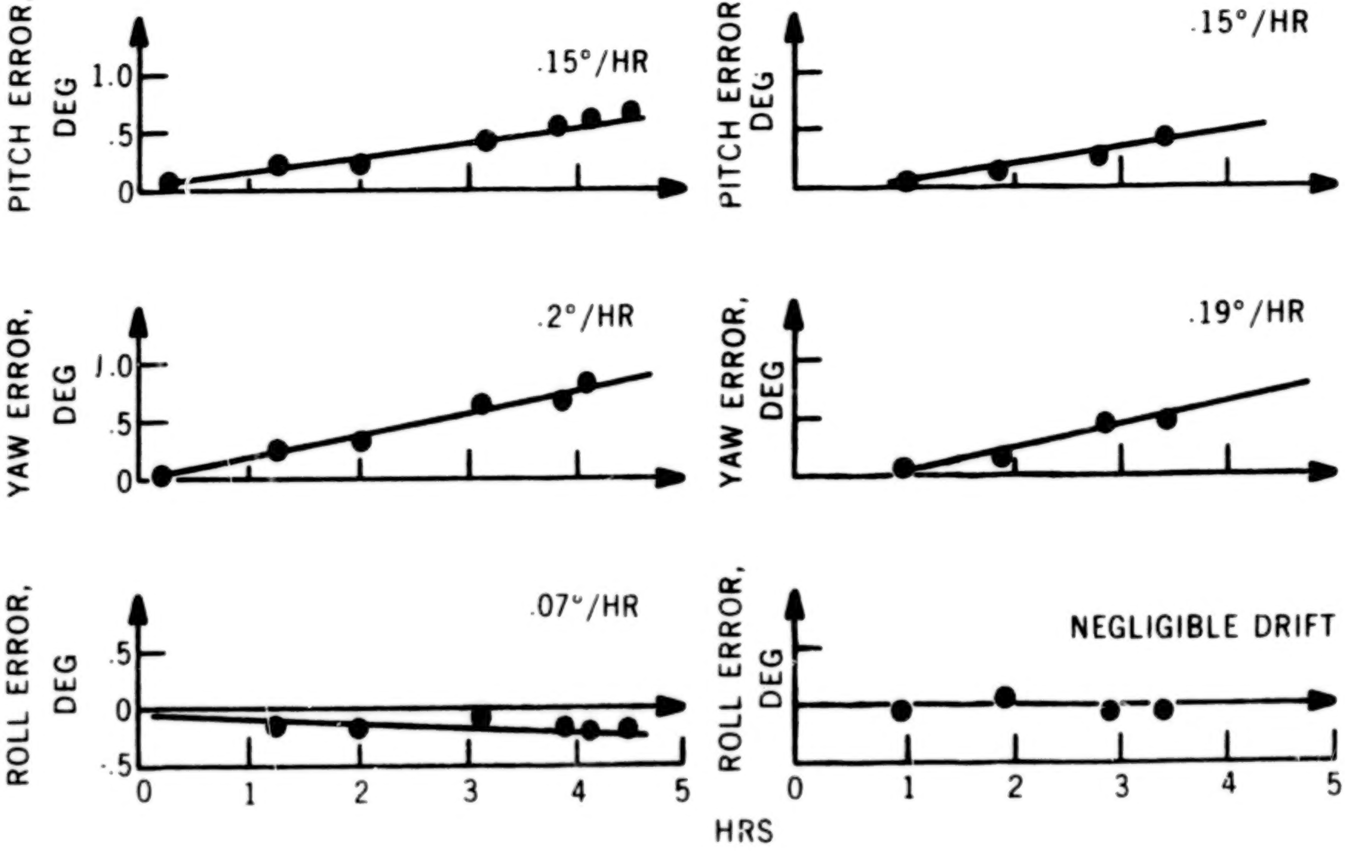


Figure 16.- Typical RAPS drift measurements

the three phase ac current to the gyros had increased by 40 percent. It was concluded that the gyro had failed.

The gyro system was turned-off and has not been used since the failure to control the spacecraft. Periodically the system is turned-on but the gyro has not recovered. This failure has put added pressure on the ground operations, but 1000 orbits has been surpassed using trackers alone without losing stabilization. If sun bathing is required, initial stabilization would be commanded and a slew to sun bathing would be executed.

Roll Search

Although the OAO-A2 did not use roll search mode for initial star acquisition, it was performed and satisfactorily achieved after the RAPS pitch gyro failure. The same equipment is used for the roll search as for initial stabilization. Fine solar acquisition is complete when the pitch rate and attitude null signals, the yaw rate and attitude null signals, and the disable eye signal exist. The programmer processes these signals through an "And" gate whose output initiates the roll search. Roll search signals are sent to SSP, the roll FWJC. The roll axis of the Observatory remains stabilized to within 0.25 deg of the sun line through use of the fine sun sensors and fine wheels. The search signal to the roll SSP causes a biased roll rate signal to be introduced, which is then fed to the roll FWJC, which in turn operates the roll high thrust jets. The search signal to the roll FWJC is used to logically inhibit the roll rate null signal which would otherwise be a false indication of rate null during roll search.

The rate of the spacecraft rotation is approximately 0.25 deg/sec, which is compatible with the dynamics of the observatory and the ability of the gimbaled star trackers to track guide stars as well as with the time of non-occultation of the sun during the orbit. The gimbal angles of each tracker are prepositioned such that at some roll attitude about the sun line each tracker will have a preselected guide star in its field of view. The gimbals remain in command mode until at least four trackers simultaneously acquire guide stars. By means of a ground command the necessary number of simultaneous star presences may be changed from four to three or to two.

When the programmed number of star trackers produce star presence signals simultaneously, the trackers are switched from command mode to track mode. The trackers will then commence to track its guide star. When the assigned trackers reduce the tracking errors to null (<2 arc minutes), the programmer removes the roll search bias. The roll gyro senses the rate and fires a roll jet to reduce the rate to zero. When the roll rate null is sensed by the programmer, the roll fine wheel is enabled and the programmer switches the Star Tracker Signal Processor (STSP) pitch and yaw star error signals into the appropriate fine wheel channels. The roll fine wheel channel is already connected to its star signal. The signal flow (Figure 17) in the control loop is then from the Gimbal Star Trackers (GST) and the Digital Logic Unit (DLU) to the STSP to the Boresight Star Tracker Logic Unit (BSTLU) to the FW&JC to the fine wheels. From the time the tracker is commanded to track mode, the tracker develops gimbal angle signals and the DLU develops gimbal error signals which are continuously processed by the STSP into control axes analog error signals.

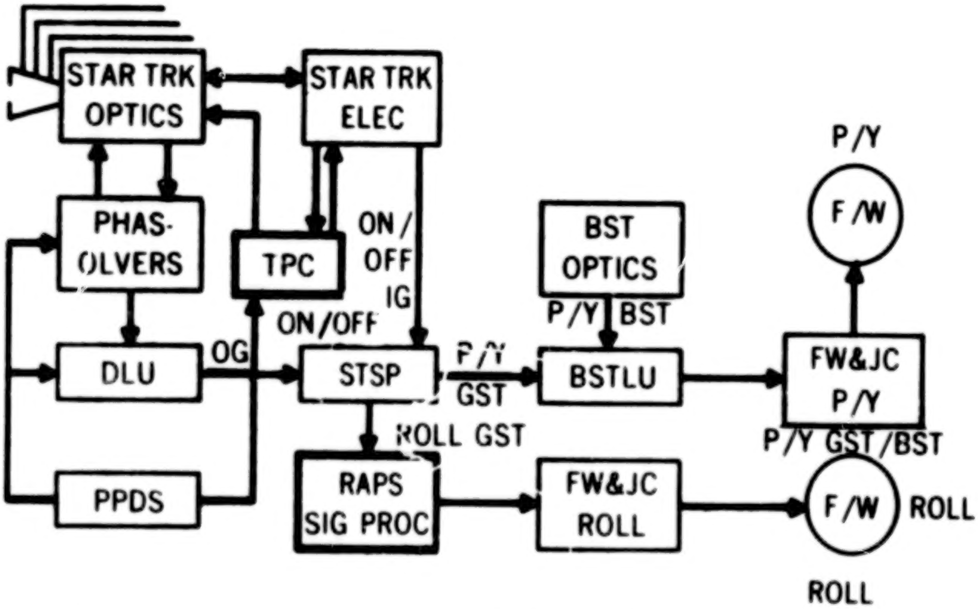

STAR TRK OPTICS
STAR TRK ELEC
P/Y
F/W
PHAS-OLVERS
TPC
ON/OFF IG
BST OPTICS
ON/OFF
P/Y BST
OG
DLU
STSP
P/Y GST
BSTLU
FW&JC P/Y
ROLL GST
P/Y GST/BST
PPDS
RAPS SIG PROC
FW&JC ROLL
F/W
ROLL
ROLL

Figure 17.- Coarse (and BST) attitude hold

Star Search Mode

Guide star acquisition by trackers was originally designed to be achieved through the roll search program about the sun line. However, this method has several basic disadvantages:

1. The attitude of vehicle at beta zero ($-X_C$ pointed to the sun) is a power negative condition.
2. Star re-acquisition is time consuming and uses large quantities of gas since initial stabilization must be commanded.
3. The availability of three stars is almost mandatory for roll search since false star presences due to earth albedo generate false star patterns when a two-star requirement is selected.

The hold-on-wheels mode, as described in the sun-bathing section of this paper, is the mode established to perform star search. In the hold-on-wheels mode, the RAPS gyros are in attitude mode and are feeding their outputs through the FW&JC to the fine wheels. The spacecraft, under gyro control, is drifting at a low rate of approximately 0.2 deg/hr on all three axes. The attitude of the spacecraft in this mode is determined by ground processing the Adcole aspect sun sensor and magnetometer data. There are four aspect sun sensors mounted on the vehicle which give (±0.5 deg) solar aspect data at any spacecraft attitude. The magnetometer data yield the attitude about the third axis.

The star search routine is as follows:

1. The system is placed in RAPS hold-on-wheels.
2. The Adcole sun sensor data and magnetometer data are collected via a ground station contact and transmitted to the Goddard Control Center where a Support Computer Program System (SCPS) performs a weighted least squares estimate of attitude.
3. The SCPS generates a search routine which commands the gimbals of the assigned tracker through a fix pattern (≈3 deg x 3 deg). When the programmed tracker acquires the guide star, it will switch from command mode to track mode and track the star.
4. After the search routine is completed and the stars acquired, coarse pointing is commanded and stellar reference is achieved.

This program has been accomplished 29 times to date on OAO-A2. It has saved time and gas and is now considered the prime mode of acquiring coarse star pointing mode.

Coarse Star Tracker Control

Coarse tracker control of the Observatory's attitude involves use of the six gimbaled star trackers as the control sensors. Only two trackers are required for three axis control. In this mode of operation, the S&C system utilizes the equipment as shown in Figure 17. The mode consists of pointing the Observatory's optical axis (roll axis) to any desired direction with a spec accuracy of 1 arc minute (1σ) and, for a period of 50 min, maintaining a spec stability of 15 arc seconds when exposed to an external disturbance torque of up to 1500 dyn-cm. The reference for determining attitude is a set of 34 selected guide stars (see Table I). Prior to the roll search, the GST's receive

TABLE 1.- OAO-A2 GUIDE STAR LIST
(Photomultiplier 1P21-S-4 Surface)

SCPS No	Name	B.S. No	G.C. No	V	U-V	B-V	S4
1	αCMa	2491	8833	-1.45	-0.03	-0.01	-1.45
2	αCar	2326	8302	-0.71	0.28	0.18	-0.50
3	αLyr	7001	25466	0	0.03	0	0
4	βOri	1713	6410	0.15	-0.66	-0.03	0.15
5	αEri	472	1979	0.49	-0.86	-0.19	0.40
6	βCen	5267	18971	0.61	-1.22	-0.24	0.42
7	αCen	5459/6	19728	-0.26	1.22	0.69	0.30
8	αAur	1708	6427	0.26	1.28	0.81	0.70
9	αVir	5056	18144	0.96	-1.18	-0.25	0.75
10	αCMi	2943	10277	0.35	0.47	0.43	0.75
11	αAql	7557	27470	0.74	0.31	0.23	1.00
12	αLeo	3982	13926	1.35	-0.47	-0.12	1.35
13	εCMa	2618	9188	1.50	-1.13	-0.21	1.35
14	αCyg	7924	28846	1.25	-0.12	0.09	1.40
15	αOri	1790	6668	1.63	1.09	-0.21	1.50
16	αGem	2890/1	10120	1.58	0.04	0.03	1.60
17	εUMa	4905	17518	1.78	-0.02	-0.03	1.80

TABLE I.-(Concluded)

SCPS No	Name	B.S. No	G.C. No	V	U-V	B-V	S4
18	βTau	1791	6681	1.66	-0.62	-0.13	1.60
19	βCMa	2294	8223	1.98	-1.20	-0.23	1.80
20	αCru	4730/1	16952	0.81	-1.23	-0.27	0.60
21	βCar	3685	12764	1.67	0.04	0.01	1.70
22	ηUMa	5191	18643	1.86	-0.85	-0.18	1.75
23	γVel	3207	11105	1.82	-1.19	-0.26	1.65
24	εSgr	6879	25100	1.84	-0.16	-0.03	1.80
25	αPav	7790	28374	1.93	-0.90	-0.20	1.80
26	χOri	2004	7264	2.06	-1.18	-0.18	1.90
27	αGem	2421	8633	1.91	0.07	0	1.90
28	σSgr	7121	25941	2.09	-0.93	-0.20	1.95
29	αGru	8425	30942	1.73	-0.59	-0.13	1.70
30	αAnd	15	127	2.06	-0.58	-0.11	1.95
31	αBoo	5340	19242	-0.06	2.52	1.24	0.80
32	αPsa	8728	32000	1.15	0.17	0.09	1.30
33	βCru	4853	17374	1.24	-1.21	-0.23	1.05
34	αTau	1457	5605	0.86	3.50	1.55	1.85

ground commands via the PPDS and DLU which define gimbal angles relative to the Observatory that correspond to the desired attitude of the Observatory reference coordinate system with respect to the celestial coordinate system. The DLU compares the commanded gimbal angles to the actual gimbal angles and derives analog error signals, which it then feeds to the STSP directly and via the star tracker gimbal resolvers. In this process, the error signals are transformed into control axes error angles and the STSP receives a pitch, yaw, and [illegible] error signal for each star tracker that is locked onto a guide star. The STSP averages the error signals on a per axis basis and provides the results as the error signals for the control loops. The existence of the combined null signal in the STSP signifies that the S&C system is maintaining the control axes error signals to less than one minute of arc (nominal), as sensed at the STSP output.

The heart of the coarse pointing loop is the star trackers and DLU. A brief description of both units is given below.

Gimbal star tracker (GST).- The star tracker consists of an integral telescope with a 1-deg field of view mounted on a 2-degree-of-freedom gimbaling system. The telescope optical system employs a 3.5-in. diameter beryllium parabola as the primary mirror with a focal length of 5 in. The secondary optical elements consist of two beryllium mirrors which separate the star light into two beams, one for each axis. The light beams are modulated by means of two apertured vibrating reeds located in the focal planes of the objective mirror at a 90-deg angle to one another. To eliminate possible electrical ambiguity, one reed vibrates at 350 cps while the other vibrates at 450 cps. After modulation, the light beams are recombined by the optic and impinged on a 1P21 photomultiplier. The gimbal system allows a ±43-deg excursion about the outer and inner gimbal axes. The tracker can recognize 2.0 magnitude or brighter stars (Table I) and track with a two gimbal accuracy of 30 arc seconds. The side trackers are equipped with a sun baffle which allows tracking to within 33 deg of the sun and 15 deg of the sunlit earth. The forward and aft trackers have short shades which have tracking angles of 63 deg to the sun and 30 deg to the sunlit earth. The tracker contains a protective shutter which closes when activated by an earth or sun sensor mounted on the shade. The sun sensor on the side trackers activates at a sun angle of 30 deg ±3 deg to the optical axis and the earth sensor operates at 8 deg ±4 deg to an earth albedo of 0.3. The shutter activates at a sun angle of 60 deg ±3 deg and earth angle of 13 deg ±4 deg on the short shade.

The tracker uses a phasolver as the gimbal readout device. It is a capacitance analog-type encoder. The phasolver is a multi-pole shaft angle to phase angle transducer whose capacitance varies with shaft angle. The reference phase signal and the signal which has been phase-shafted by the phasolver disks are detected and are used to open and close gates of a digital counter. The phasolver has a resolution of 5 arc sec.

Digital logic unit (DLU).- The DLU is a time-shared device which receives signals from 12 star tracker gimbal pick-offs (phasolvers) which indicate gimbal angles, sequentially compares the magnitude of each gimbal angle with a commanded angle, and developes as outputs 12 analog error signals, proportional to the difference between the gimbal angles and the gimbal commands. Holding circuits maintain the value of the error signals essentially constant between read-out periods of each gimbal. The binary command signals and various timing signals are received from an external clock system (PPDS).

Boresighted star tracker mode.- The boresighted star tracker mode (Figure 17) is considered part of coarse pointing and involves use of the Boresighted Star Tracker (BST) to provide attitude error signals for the pitch and yaw control loops. The roll error signal is provided by the GST's. The mode consists of orienting and holding the Observatory's pitch and yaw axes to within 20 arc seconds for a sixth magnitude star. The roll attitude is held to within one minute of arc using the GST's. The BST is aligned to the experiment and is used as a calibration tool to correct misalignments. Initially, the Observatory is slewed by ground command to that attitude where its optical axis ($+X_C$) is pointing at a guide star selected for the BST. After slewing, the spacecraft is stabilized in the coarse pointing mode which results in pointing of the optical axis to within one arc minute of the desired direction. The BST has a 5-arc minute half-cone angle, instantaneous field of view, which can be electronically offset by ground commands in the range of ±90 min of arc about the original zero in both pitch and yaw. When the BST acquires the star within its field of view, it sends a star presence signal to the BSTLU, which processes it with the combined star tracker null signal received from the STSP. When both signals are present, the BSTLU switches within itself from the GST to the BST pitch and yaw error signals, which in turn are sent to the pitch and yaw fine wheel control loops.

Coarse pointing - flight performance.- The coarse pointing mode met the 2-arc minute and 15-arc second jitter requirements as evidenced by Figures 18 and 19. Figure 18 is a plot of GST gimbal error data vs. time. The BST was tracking a star during this period, but the GST was controlling the spacecraft. The same conditions existed for the Figure 19 data. The BST data indicate that the apparent spacecraft motion is less than 3 arc seconds. These data include BST noise, which was not separate

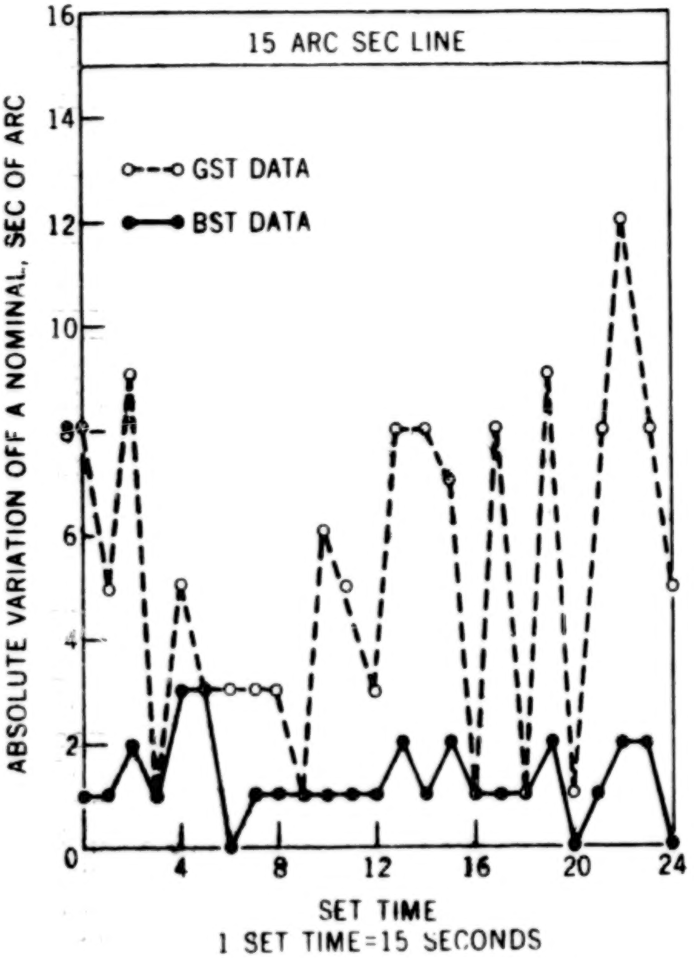


Figure 18.- Typical spacecraft holding variation

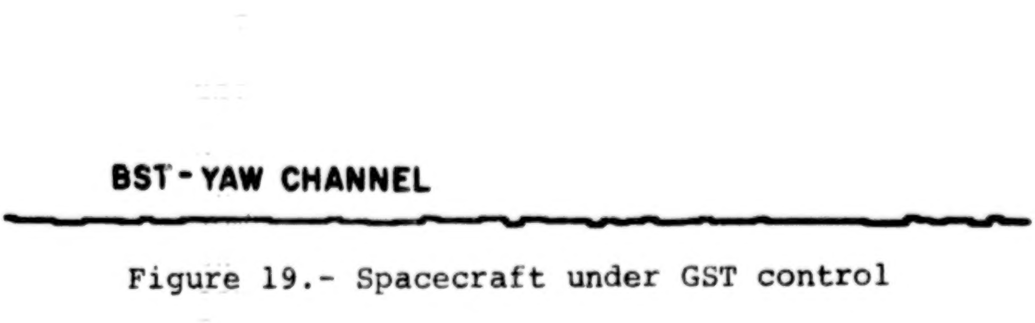


Figure 19.- Spacecraft under GST control

from the data. The spacecraft motion is felt to be less than 3 arc seconds since the wheel speed of the momentum wheels was less than 5 percent at this time with a 1 percent variation. The gain of the coarse loop is 26 Vac to the fine wheel motor for 20-arc second STSP error. Therefore, a spacecraft steady error of 1 arc second with a fraction of 1-arc second jitter is probable.

Figure 19 shows a dc shift in BST error when a gimbaled tracker is brought into or removed from the star tracker loop control. This shift, which is approximately 30 to 45 arc seconds, is mainly due to tracker misalignment errors. A star tracker misalignment calibration procedure was conducted in space to reduce the errors to this magnitude.

Star tracker misalignment calibration.- To meet the 1-arc minute pointing requirement under star tracker control, an in-orbit misalignment calibration of the star trackers was required. Launch vibration and thermal shifts caused the trackers to shift in the order of 3 arc minutes in both inner and outer gimbals. The calibration procedure to align the trackers is given in detail in reference 1. Since the trackers are two gimbal devices, there are four misalignment parameters:

1. Rotation about the zenith position
2. Rotation about the inner gimbal
3. Rotation about the outer gimbal
4. Inner gimbal null shift.

By using trackers in pairs, the correct angle between guide stars can be expressed in terms of measured gimbal angles and the misalignment parameters. Since the OAO-A2 operates on both sides of the spacecraft (A and B), a set of tracker misalignment numbers was acquired for operation on each side because of the bending motion of the spacecraft caused by thermal gradients.

This procedure is extremely important to assure precise pointing and will be mandatory for OAO Flight 3. OAO-A2 experiment (WEP) has ±5-arc minute circular field of view. The Goddard Experiment Package (GEP), which will fly on Flight 3, has a 2.5-arc minute field of view. The GEP will control the vehicle in the fine pointing mode.

Star tracker - flight performance.- The following major areas of tracker performance were evaluated in flight:

1. High voltage performance
2. Star magnitude recognition level
3. Servo operation.

High voltage performance: One of the major problems which occurred on OAO Flight A-1 was tracker high voltage arcing in orbit. An ambitious program was initiated after this failure to redesign the tracker high voltage section to operate from sea level pressure to 10^{-6} mm Hg without high voltage arcing. To make the 900-Vdc system arc proof, the following changes to the tracker were made:

1. The base of the photomultiplier tube housing the dynode resistors was potted.
2. The exposed terminals on the high voltage capacitor module were potted.
3. The high voltage transformer was potted.
4. A current limiting resistor was added in series with the anode lead.
5. New high voltage wire was used.

All trackers were tested on the ground through the partial pressure region with no evidence of arcing.

The tracker high voltage was turned-on in Orbits 26 to 28. The two days of vehicle outgassing was considered adequate to assure hard vacuum environment before turn-on. No evidence of arcing or corona was experienced during turn-on or to date.

Star magnitude recognition level: A standard procedure for star simulator calibration and photomultiplier tube spectral response compensation was developed by GSFC for tracker calibration. This procedure involves a light source which is used as a standard of irradiance and computer programs which relate this standard source to real stars. A detailed description of the calibration is contained in reference 2.

Flight data indicated that the guide stars always appeared brighter than anticipated and by varying amounts from star to star. The deltas, from star to star, vary from $0\overset{m}{.}16$ to $0\overset{m}{.}55$. Table II shows this variation.

Another observation which is not fully understood is the apparent loss of sensitivity of the trackers over a period of time. The sensitivity dropped off rapidly in the first 200 orbits ($0\overset{m}{.}2$) and then a gradual decline of $0\overset{m}{.}05$/month has been observed.

The main cause of the decrease in sensitivity is felt to be the exposure of the photomultiplier to the bright earth before the protective shutter closed. A cadmium sulfide earth sensor is mounted on the shade of each tracker to sense the bright earth and close a protective shutter. As the tracker impinges upon the

TABLE II.- GUIDE STAR DATA

Guide Star	Type	S-4 Mag.	ST 1	ST 2	ST 3	ST 4	ST 5	ST 6
1	A1	-1.45						
2	F0	- .50						
3	A0	0		-.16	-.10	-.10	-.12	-.27
4	B8	.15	-.10			-.10		-.40
5	B5	.40	-.05					
6	B1	.42						-.05
7	B2	.30		.10	0	.12		-.10
8	G8	.70	.37		.27		.75	
9	B1	.75		.50			.25	.25
10	F5	.75	.37		.35			
11	A7	1.00			.55	.75		
12	B7	1.35	.99		.95	1.00	1.00	
13	B2	1.35	1.00		.95		1.05	
14	A2	1.40	1.13	1.00		1.20		
15	B5	1.50						
16	A1	1.60			1.20	1.28		1.38
17	A0	1.80			1.45		1.60	1.47
18	B7	1.60						

TABLE II.- (Concluded)

Guide Star	Type	S-4 Mag.	ST 1	ST 2	ST 3	ST 4	ST 5	ST 6
19	B1	1.80			1.43	1.60		
20	B1	.60	.25					
21	A0	1.70	1.45	1.48				
22	B3	1.75		1.22	1.30			1.37
23	WC7	1.65	1.15	1.10	1.10			
24	B9	1.80						
25	B3	1.80			1.26	1.45	1.50	1.37
26	B5	1.90	1.60		1.52			
27	A0	1.90						
28	B2	1.95						
29	B5	1.70			1.33	1.40	1.42	1.33
30	B9	1.95	1.84	1.55			1.80	
31	K2	.80		.50	.50		.35	.40
32	A3	1.30		.86	1.00	1.03		
33	B0	1.05	.70	.55				
34	K5	1.85	1.55			1.62	1.62	1.42

earth, the shutter closes. However, the sensor is set to actuate at 8 deg ± 4 deg on an earth albedo of 0.3. The high voltage telemetry indicates that significant anode currents (1.6 μa) are drawn before the shutter closes possibly causing degrading of the PMT. Tests conducted at the vendor have been inconclusive.

Servo operation: The tracker servos were first checked in Orbit 1 when the trackers were decaged. All trackers performed normally.

The amount of overshoot of the gimbaled star trackers when commanded to new gimbal angles in a weightless environment is of interest. To determine the amount of overshoot the following commands were issued to the trackers:

Max. Overshoot (Rounded Off to Nearest Minute or Degree)

Gimbal Angle Change	ST 4 Tel. Temp = -8°		ST 1 Tel. Temp = -18°		ST 2 Tel. Temp = 37°	
	IG	OG	IG	OG	IG	OG
5°	8'	6'	12'	6'	5'	11'
10°	13'	40'	14'	53'	6'	13'
20°	2°	3°	2°	4°	1°	3°
35°	7°	10°	6°	11°	6°	8°
IG 50°	13°	--	13°	--	14°	--
OG 45°	--	15°	--	17°	--	14°

It can be concluded that the average overshoot is approximately 30 percent of the change in commanded angles of 30 deg or more. Also, from the data available, colder trackers tend to overshoot more and take longer to settle out than warmer trackers. This agrees with ground testing of trackers.

Tracker anomalies.- The trackers have exhibited several anomalies:

1. Tracking bright objects
2. Gimbal binding
3. Outer gimbal command failure - tracker 3.

Tracking bright objects: There has been a number of light sources tracked by the trackers in addition to guide stars, which have caused loss of coarse pointing. Several of the false stars were determined to be the Centaur vehicle. However, the majority of the false stars has not been explained. There is approximately

1700 pieces of space "junk" in orbit which make many false stars. An inhibit circuit, which removes a tracker from the control loop if the gimbal error exceeded 42 arc minutes, is presently enabled. This circuitry aids greatly in maintaining tracker control; however, it is not foolproof, since a one star tracker control situation can result which is unstable. The OAO has recently passed the 1000 orbit mark without losing control.

Gimbal binding: Three trackers have had gimbal-binding problems to date. Fortunately, each stuck gimbal has broken free upon exercising it and started working again. The dimensional tolerances on thermal covers, cable bundles, cable ties are very tight in the area of the tracker gimbals. Power scans of the gimbal position are made periodically as part of spacecraft checkout, but no precise cause has been established for the in-flight binding troubles. Gimbal bearings are presently being examined for galling or lubrication problems, but results to date are negative.

Tracker failure: Star Tracker 3 outer gimbal has become inoperative rendering the tracker useless. A line has opened between the phasolver and the DLU. The exact cause has not been determined to date. This reduces the available trackers to five, which puts added pressure on ground operations.

Bore sight tracker - flight performance.- The bore sight tracker has been used to align the gimbal star trackers to the experiment. Since the BST has offset capability, it can zero out any misalignment between it and the experimental optical axis.

The BST can hold the spacecraft in pitch and yaw to 20 arc seconds when tracking a sixth magnitude star. Although the BST has been used to control the spacecraft, it is not part of the SCPS program. Since the BST can track 2000 stars, its star catalog was not made part of the ground system program. Table III lists the stars the BST has acquired.

The BST was not designed to operate over the partial pressures. The BST was turned ON during the fourth day of flight operations with no evidence of high voltage arcing or corona.

Re-orientation or Slewing Mode

Slewing of the spacecraft from one orientation to another is accomplished by use of the equipment shown in Figure 20. Successive rotations are applied about each control axis by commanding a precise number of revolutions of each coarse momentum wheel in sequence. The stabilized coarse pointing mode serves as the

TABLE III.- STARS ACQUIRED BY BST

Santiago	62	Beta Carinae	1.67
Santiago	63	Beta Carinae	1.67
Santiago	64	Beta Carinae	1.67
Orroral	69	Beta Carinae	1.67
Orroral	71	Gamma Vela	1.82
Rosman	568	Psi Eridanus	4.79
Madgar	568	Psi Eridanus	4.79
Madgar	570	Beta Orion	0.15
Santiago	571	Beta Orion	0.15
Santiago	572	Beta Orion	0.15
Santiago	574	Epsilon Lepus	3.19
Orroral	576	Epsilon Lepus	3.19
Quito	577	S50896	6.9
Rosman	847	Eta Ursa Major	1.86
Orroral	847	Eta Ursa Major	1.86
Santiago	981	Sirius	-1.5
Rosman	1204	Beta Orion	0.1
Santiago	1206	Mu Lepus	3.3
Rosman	1218	Upsilon Orion	4.6
Rosman	1230	Theta^2 Orion	5.1
Rosman	1231	42 Orion	4.6
Madgar	1231	HD37-356	6.2
Santiago	1249	Delta Orion	2.2
Santiago	1252	Beta Eridanus	2.79
Orroral	2156	S36285	6.35

TABLE III.-(Continued)

Madgar	1257	S36285	6.35
Rosman	1258	S36430	6.24
Madgar	1258	S36430	6.24
Santiago	1265	S35299	5.7
Orroral	1270	Omega Orion	4.52
Rosman	1272	Alpha Orion	0.8
Madgar	1272	Alpha Orion	0.8
Rosman	1273	Alpha Orion	0.8
Madgar	1273	Alpha Orion	0.8
Quito	1276	S34989	5.8
Santiago	1279	S34989	5.8
Madgar	1288	S35600	5.7
Madgar	1289	S35600	5.7
Santiago	1294	Theta Auriga	2.65
Orroral	1296	Alpha Auriga	0.1
Orroral	1297	Alpha Auriga	0.1
Rosman	1299	Alpha Auriga	0.1
Rosman	1301	Alpha Auriga	0.1
Madgar	1602	Mars	
Rosman	1610	Tau Scorpius	2.8
Rosman	1652	Mu Serpens	3.53
Rosman	1653	Mu Serpens	3.53
Rosman	1667	Gamma Hercules	3.76
Orroral	1668	Gamma Hercules	3.76
Rosman	1682	Alpha Crab	2.27

TABLE III.-(Continued)

Orroral	1682	Alpha Crab	2.27
Rosman	1684	Alpha Bootes	.06
Madgar	1686	Zeta Bootes	3.78
Quito	1692	Alpha Serpens	2.65
Rosman	1695	Epsilon Serpens	3.7
Rosman	1696	Epsilon Serpens	3.7
Quito	1706	Zeta Ophiuchus	2.57
Rosman	1708	142883	5.84
Quito	1710	142184	5.4
Rosman	1711	142184	5.4
Rosman	1712	142165	5.41
Rosman	1713	142165	5.41
Madgar	1714	142290	5.44
Madgar	1715	13 Scorpius	4.58
Rosman	1723	Mars	
Rosman	1725	Chi Ophiuchus	4.43
Madgar	1729	Chi Lupus	3.94
Rosman	1739	Mu Norma	4.90
Santiago	1743	Mu 1 Scorpius	3.02
Santiago	1744	148688	5.32
Orroral	1748	152236	4.74
Quito	1748	152236	4.74
Santiago	1761	150989	5.55
Orroral	1777	Pi Lupus	3.88
Madgar	1785	135591	5.43

TABLE III.-(Concluded)

Orroral	1791	Alpha Lupus	2.29
Quito	1791	Alpha Lupus	2.29
Rosman	1793	Eta Centaurus	2.30
Madgar	1794	Eta Centaurus	2.30
Rosman	1796	Psi Centaurus	4.04
Quito	1798	Theta Centaurus	2.05
Quito	1799	Theta Centaurus	2.05
Santiago	1803	Beta Centaurus	.61
Quito	1804	Beta Centaurus	.61
Santiago	2325	Alpha Virgo	.96
Orroral	2330	3C273	
Santiago	2336	Beta Leo	2.13
Rosman	3094	Beta Crux	1.26
Rosman	3109	Gamma Vela	1.82
Rosman	3110	Gamma Vela	1.82
Rosman	3111	Gamma Vela	1.82
Quito	3423	Eta Auriga	3.2
Santiago	3424	Eta Auriga	3.2
Santiago	3426	Beta Tau	1.9

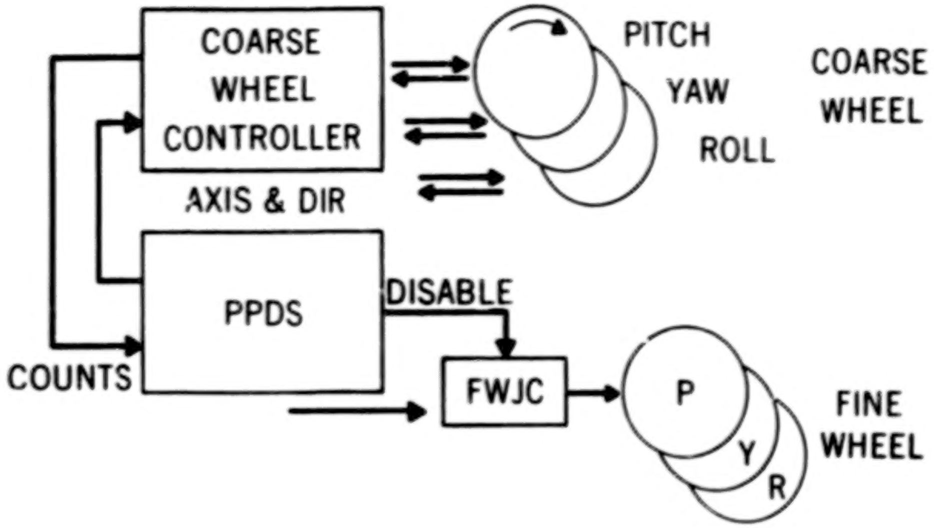


Figure 20.- Re-orientation

initial attitude reference from which the required number of revolutions of each wheel is computed on the ground. The Observatory rotates through a precise angle and is caused to stop by braking the wheel to an absolute stop. The star trackers track their guide stars during this operation. Prior to slew, those star trackers that would approach their gimbal limits are commanded to transfer to another preselected guide star. The star tracker gimbal angles corresponding to the Observatory attitude at the end of each slew are computed on the ground and transmitted to the Observatory. Following each axis of slew the control system restabilized in the coarse pointing mode by comparing the actual star tracker gimbal angles to the commanded values to derive pitch, yaw, and roll attitude error signals. The coarse wheels are capable of open loop slewing of the Observatory through 2 deg in 40 sec and 30 deg in 250 sec for 1800-slug/ft^2 vehicle. During slews, the fine wheels are disabled. The wheels are unloaded before slewing to reduce the momentum added to the vehicle due to wheel rundown. The unloading of the fine wheels is accomplished by two methods (Figure 21):

1. Pneumatic unloading
2. Magnetic unloading.

In pneumatic unloading,the fine wheels are unloaded automatically by the low-thrust gas-reaction jet equipment if the wheel speed exceeds 75 percent of no-load speed for all modes except roll search where automatic unloading occurs at 40 percent of no-load speed. The ground commands, permission to unload - 5 percent and permission to unload - 40 percent, cause the fine wheels to be unloaded if the wheel has reached or exceeded 5 or 40 percent of no-load speed at the time the command is executed. The low thrust jets

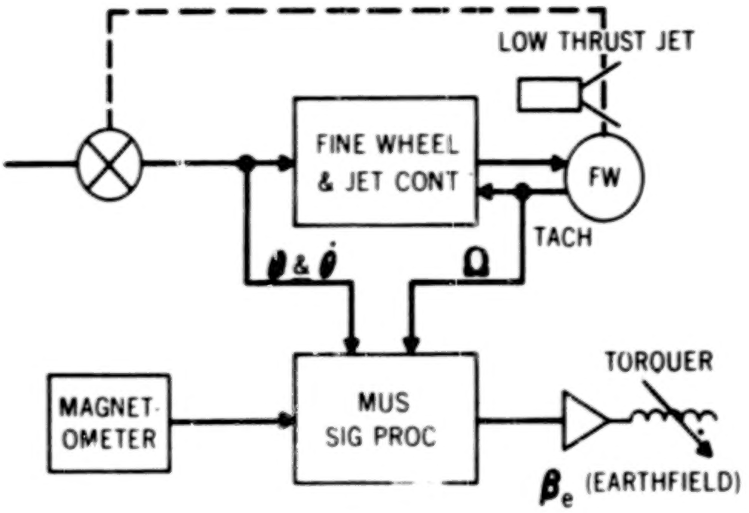


Figure 21.- Momentum dumping

accomplish unloading by applying torques on the Observatory of sufficient magnitude and appropriate direction to reduce the fine wheel speeds to less than 8 percent of the no-load speed.

In magnetic unloading, the fine wheels are unloaded continuously by the magnetic unloading equipment as a function of wheel speed and attitude error. The MUS consists of three basic components:

1. Three magnetometers - sense a component of earth's field.
2. Signal processor - performs algebraic summation of magnetometer outputs and the FW&JC signals.
3. Three torquer bars - generate the magnetic field proportional to drive currents.

The magnetic unloading equipment receives signals from the FWJC's and magnetometers which are processed into drive signals for its magnetic torquer coils. The fields produced by the torquer coils interact with the earth's magnetic field to produce the torque necessary for unloading. The system can generate 10,000 cgs units, which is equivalent to 3000 dyn-cm of torque when reacting with an earth field of 0.3 gauss at 500 miles.

Re-orientation mode - flight performance.- Figure 22 gives a comparison of predicted and actual slew angles versus spacecraft slew time plus settling time. The curves show that the settling time can be reduced by turning the MUS off during slew. This fact exists since, when the fine wheels are disabled, the FW&JC signal to the MUS is opened circuited and the MUS saturated generates a disturbance torque which fights the slew. This causes

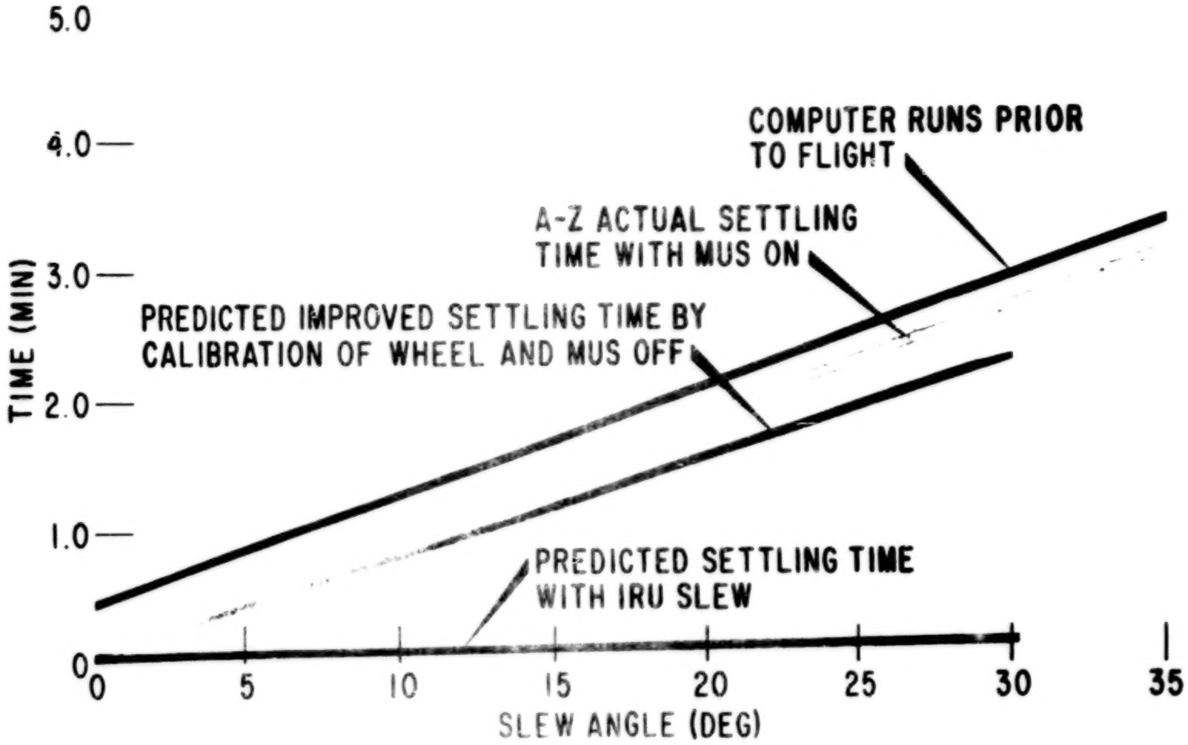


Figure 22.- Open loop settling time curve (12½% wheel speeds)

a larger angle error after slew to settle from under tracker control. The coarse wheel torque and no-load wheel speed have been termed to be nominal.

The predicted settling time with the Inertial Reference Unit (IRU) controlling the spacecraft during a close loop slew is shown to emphasize the vast improved settling time that will be realized on OAO Flight 3.

MUS - flight performance.- The magnetic unloading system has kept the fine wheel speeds below 20 percent when it is operated. The OAO slew sequence contained 5 percent unload commands on all three axes before slew. This requirement was necessary to minimize errors at the end of slew due to fine wheel rundown during the slew. With the MUS ON, the requirement is deleted since the wheel speeds are kept low throughout the orbit.

The disturbance torques observed at the 500-mile orbit due to magnetic, gravity gradient, solar pressures, and drag are approximately 700 to 1000 dyn-cm (average). The MUS has a 3000-dyn-cm torque capability.

Jet unloading performance.- The low thrust jets unload the wheels when the wheel speed exceeds 75 percent. The control torque of the pitch and yaw jets is 0.0089 ft-lb and roll torque is

.0064 ft-lb. Flight data indicate that pitch and yaw unloading torques are nominal; however, the roll torque is reduced by 50 percent. After extensive analysis, it was determined that the roll jet is impinging on the solar paddle causing the reduction in torque. The problem is not severe since the MUS system has eliminated the need for gas unloading.

Pneumatics.- The pneumatics system consists of eight tanks, each holding 8 lb of dry nitrogen at a pressure of 3250 psi. The tanks are connected through a common manifold and are filled through a single check valve. There is a primary and secondary system with individual control through two latch valves. The high thrust jets are redundant, while there is only one set of low thrust jets. The 3250 psi pressure is lowered to 38 psi through a pressure regulator. The low thrust system is regulated to 5 psi. A schematic of the system is shown in Figure 23. The pneumatic system has operated normally through the flight except for the roll jet impingement problem. The system is presently setting at 1500 psi with 34 lb of gas.

Fine Pointing Mode

The fine pointing mode is defined as switching from tracker control to an optical experiment sensor and controlling the spacecraft pitch and yaw axis by the sensor outputs. The Wisconsin Experiment Package (WEP) does not have a fine guidance sensor since the coarse pointing accuracy of 1 arc minute was adequate. The GEP experiment will point the pitch and yaw axes to within 1 arc second for bright stars. The Princeton Experiment Package (PEP) will point the Flight 4 OAO vehicle to within 0.1 arc second.

ACKNOWLEDGEMENT

Without the Goddard-Grumman Aerospace team which integrated, tested, and operates the OAO-A2 in orbit, this paper could not have been written concerning a successful OAO flight. The statement "A job well done" is extended to all the participants.

REFERENCES

1. des Jardins, R.: "In-orbit Star Tracker Misalignment Estimation on the OAO." X-542-69-418, September 1969.

2. Draper, Lawrence T.: "Star Tracker Calibration." NASA-TND-4594, June 1968.

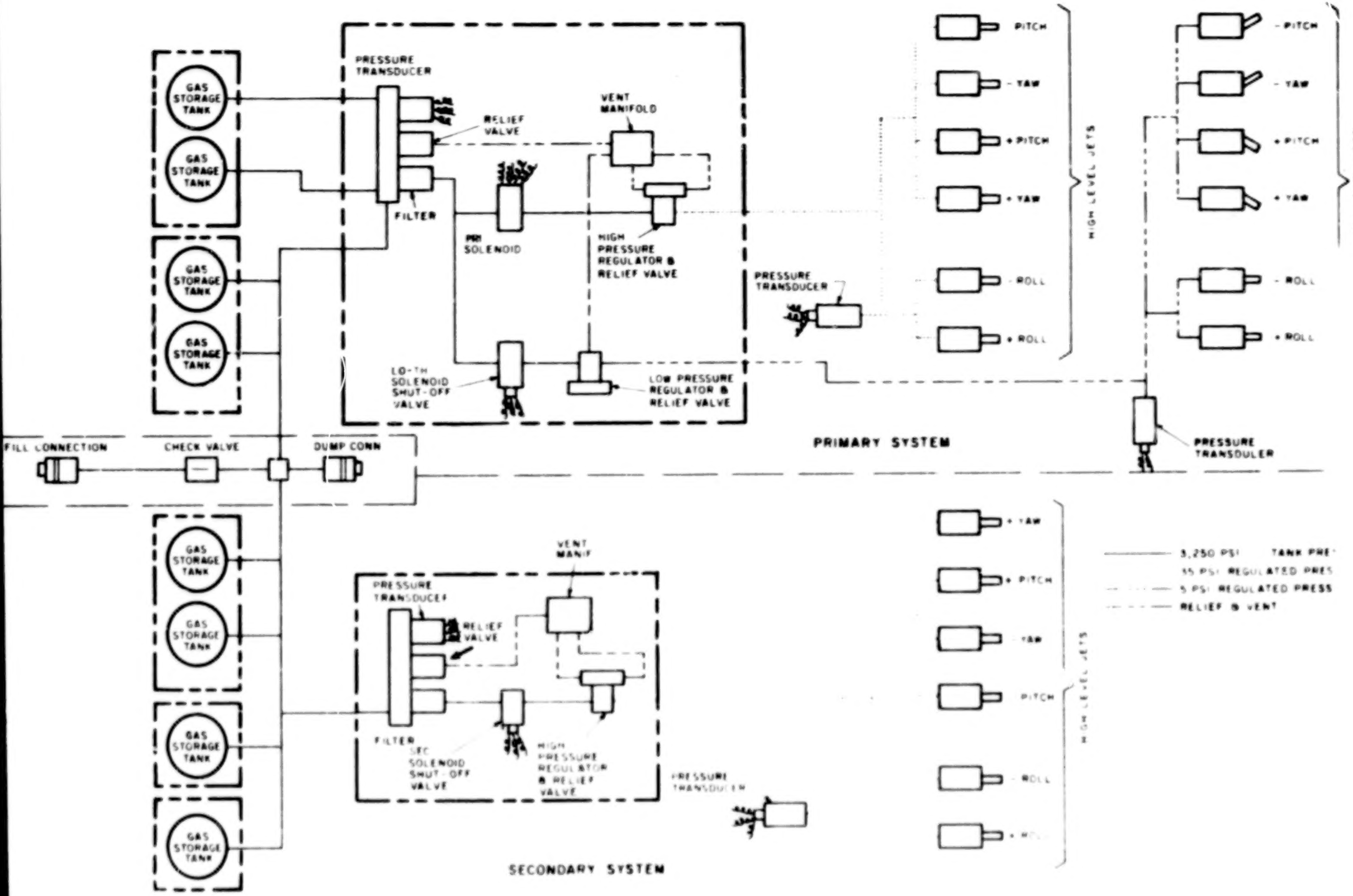

GAS STORAGE TANK
PRESSURE TRANSDUCER
RELIEF VALVE
VENT MANIFOLD
FILTER
PRI SOLENOID
HIGH PRESSURE REGULATOR & RELIEF VALVE
LO-TH SOLENOID SHUT-OFF VALVE
LOW PRESSURE REGULATOR & RELIEF VALVE
FILL CONNECTION
CHECK VALVE
DUMP CONN
PRIMARY SYSTEM
PRESSURE TRANSDUCER
PRESSURE TRANSDULER
PITCH
- YAW
+ PITCH
+ YAW
- ROLL
+ ROLL
HIGH LEVEL JETS
- PITCH
- YAW
+ PITCH
+ YAW
- ROLL
+ ROLL
LOW LEVEL JETS
+ YAW
+ PITCH
- YAW
- PITCH
- ROLL
+ ROLL
HIGH LEVEL JETS
3,250 PSI TANK PRE
35 PSI REGULATED PRES
5 PSI REGULATED PRESS
RELIEF & VENT
VENT MANIF
PRESSURE TRANSDUCER
RELIEF VALVE
FILTER
SEC SOLENOID SHUT-OFF VALVE
HIGH PRESSURE REGULATOR & RELIEF VALVE
PRESSURE TRANSDUCER
SECONDARY SYSTEM

Figure 23.- Pneumatic system

124

3. General Electric Company: "OAO Flight B Mod. Stabilization and Control Subsystem and Component Design Requirements." July 1967.

4. Grumman Aerospace Corporation: "OAO-A2 Thirty-Day Evaluation Report." April 1969.

5. Davenport, Paul: "OAO-A2 Guide Star List." Memorandum, August 1968.

BLANKPAGE

APPLICATIONS OF THE HYBRID COORDINATE METHOD TO THE TOPS AUTOPILOT

By G. E. Fleischer
Jet Propulsion Laboratory

SUMMARY

Preliminary results are presented from the application of the hybrid coordinate method to modeling TOPS (Thermoelectric Outer Planet Spacecraft) structural dynamics. Included are computer simulated responses of the vehicle which illustrate the interaction of relatively flexible appendages with an autopilot control system.

Comparisons are made between simplified single-axis models of the control loop, with spacecraft flexibility represented by hinged rigid bodies, and a very detailed three-axis spacecraft model whose flexible portions are described by modal coordinates. While single-axis system, root loci provided reasonable qualitative indications of stability margins in this case, they were quantitatively optimistic when matched against responses of the detailed model.

The hybrid coordinate approach promises to be valuable for TOPS attitude control studies, since it accommodates any type of control mechanism on the rigid portion of the vehicle, e.g. momentum wheels, provided the rigid body motion is "small" (or is a constant rate rotation) and flexible appendage deformations are linearly elastic and "small."

Computer simulations showed the necessity for at least a 2 to 1 reduction in nominal control loop gain in order to avoid instabilities, particularly in yaw, due to significant flexible appendage interaction.

TOPS AUTOPILOT

The Thermoelectric Outer Planet Spacecraft (TOPS), as presently configured (see Figure 1), consists of (1) a central compartment, containing various electronic equipment, which is assumed to be quite rigid, (2) a pair of large, relatively flexible, boom-like structures which fold out on opposite sides of the equipment compartment carrying radioisotope generators and scientific instruments, (3) a large diameter (14 ft), deployable parabolic antenna, which is also quite flexible, and (4) a pair of very flexible magnetometer booms telescoping out approximately 10 ft from the central compartment.

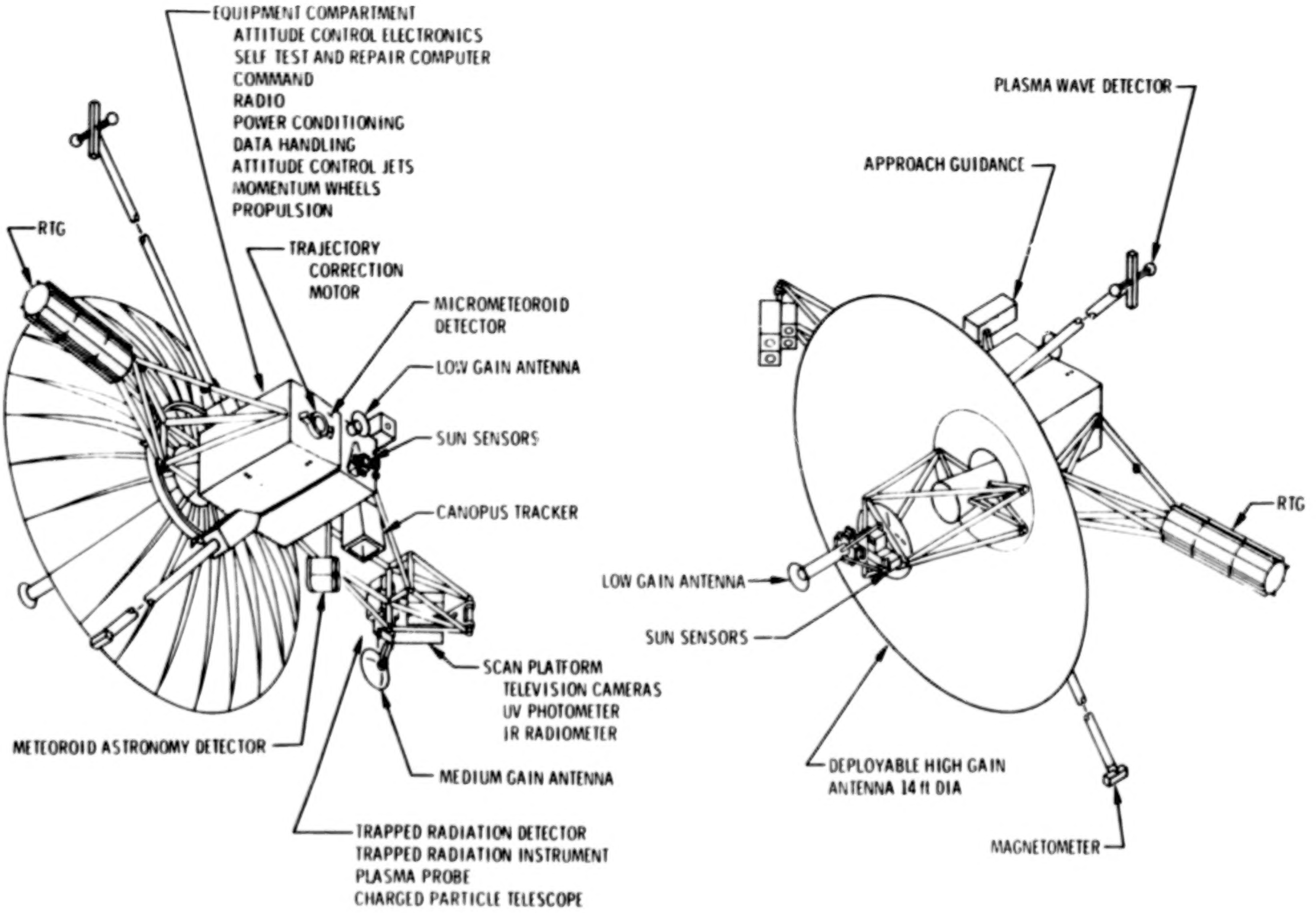

EQUIPMENT COMPARTMENT
ATTITUDE CONTROL ELECTRONICS
SELF TEST AND REPAIR COMPUTER
COMMAND
RADIO
POWER CONDITIONING
DATA HANDLING
ATTITUDE CONTROL JETS
MOMENTUM WHEELS
PROPULSION
RTG
TRAJECTORY CORRECTION MOTOR
MICROMETEOROID DETECTOR
LOW GAIN ANTENNA
SUN SENSORS
CANOPUS TRACKER
SCAN PLATFORM
TELEVISION CAMERAS
UV PHOTOMETER
IR RADIOMETER
METEOROID ASTRONOMY DETECTOR
MEDIUM GAIN ANTENNA
TRAPPED RADIATION DETECTOR
TRAPPED RADIATION INSTRUMENT
PLASMA PROBE
CHARGED PARTICLE TELESCOPE
PLASMA WAVE DETECTOR
APPROACH GUIDANCE
RTG
LOW GAIN ANTENNA
SUN SENSORS
DEPLOYABLE HIGH GAIN ANTENNA 14 ft DIA
MAGNETOMETER

Figure 1.- Thermoelectric outer planet spacecraft configuration 12J

During cruise or attitude maneuver periods, the proposed control is a 3-axis momentum wheel system using gas jets to periodically desaturate the wheels. During trajectory correction thrust periods, a 25-to 100-lb thrust gimbaled engine is used to provide attitude stability about the pitch and yaw axes, while roll disturbances, due largely to exhaust gas swirl torque, are countered by the cruise wheel system.

Initial design study phases of the TOPS mission have attempted to determine the feasibility of using a particular gimbaled-engine autopilot control system originally designed for a Mars-orbiter type spacecraft (ref. 1). Figure 2 shows a linearized and simplified single axis version of the autopilot loop. While the blocks defining gyro and gimbal servo characteristics are reasonably fixed, the autopilot compensation block (expanded in Figure 3) and total loop gain are relatively free for accommodating particular spacecraft configurations and associated performance requirements, although they may or may not be effective in this.

SIMPLIFIED DYNAMIC ANALYSIS AND ROOT LOCUS STUDIES

In order to obtain some initial qualitative indication of the effects of vehicle flexibility on control system performance, a simple analytical model was devised under the assumption that flexibility could be approximated by hinged rigid bodies. Equations of motion for such models are usually derived from the Lagrangian formulation. The energy expressions are assembled, including the effects of discrete springs and dampers at the hinges, and differentiated, producing the desired relations. However, identical equations will be derived here by using the general result of Hooker and Margulies (ref. 2) for a system of interconnected rigid bodies.

The "barycenter formulation" of Hooker-Margulies is based on the general vector equation:

$$\Phi_n^{\star} \cdot \dot{\underline{\omega}}_n + \underline{\omega}_n \times \Phi_n^{\star} \cdot \underline{\omega}_n = \underline{T}_n + \sum_{j \varepsilon J} \underline{T}_{nj} + \underline{b}_{nn} \times \underline{F}_n$$

$$+ \sum_{i \neq n} \underline{b}_{ni} \times (\underline{F}_i + m\ddot{\underline{b}}_{in}) \qquad (1)$$

where:

$$\underline{b}_{nn} = - \frac{1}{m} \sum_{i \neq n} m_i \underline{\ell}_{ni}$$

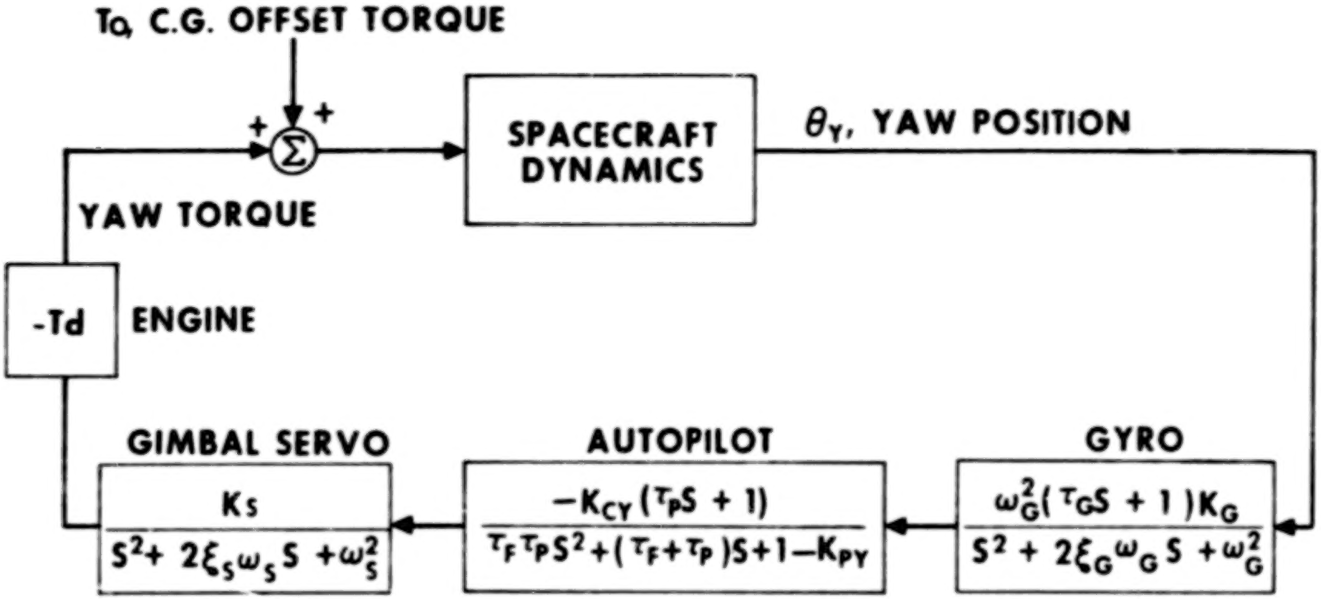

To, C.G. OFFSET TORQUE
+
+
Σ
SPACECRAFT DYNAMICS
θY, YAW POSITION
YAW TORQUE
-Td
ENGINE
GIMBAL SERVO
Ks / S²+ 2ξSωS S +ωS²
AUTOPILOT
−KCY (τPS + 1) / τFτPS²+(τF+τP)S+1−KPY
GYRO
ωG²(τGS + 1)KG / S² + 2ξGωG S +ωG²

Figure 2.- Linearized and simplified yaw axis autopilot loop

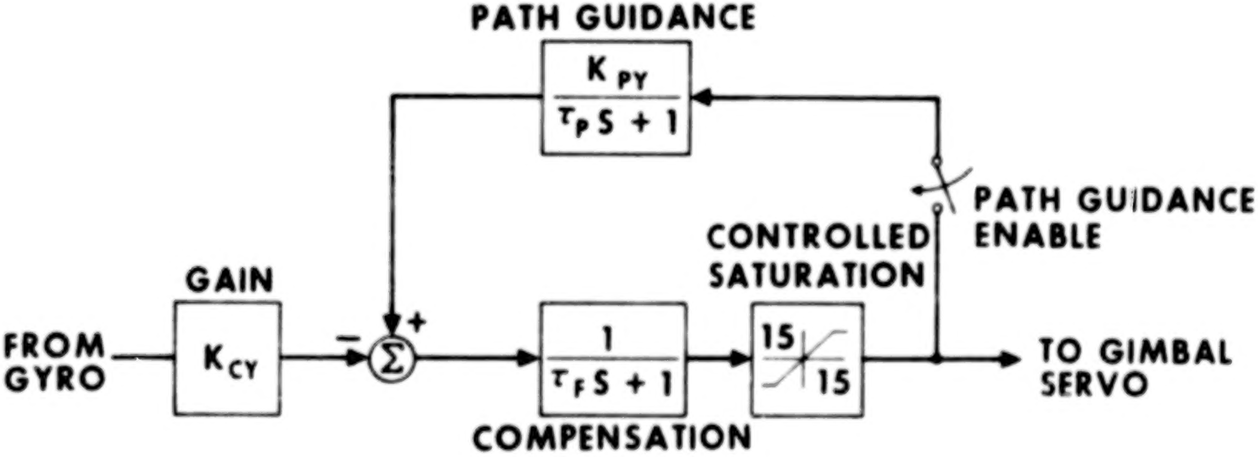

PATH GUIDANCE
K PY / τP S + 1
PATH GUIDANCE ENABLE
CONTROLLED SATURATION
GAIN
FROM GYRO
KCY
−
+
Σ
1 / τF S + 1
15
15
TO GIMBAL SERVO
COMPENSATION

Figure 3.- Autopilot block diagram

m = total mass of the system = $\sum_{i}^{N} m_i$

N = no. of rigid bodies in the system

$\underline{\ell}_{ni}$ = vector from c.m. of body n to the hinge connecting n to body i

$\underline{b}_{ni}$ = $\underline{b}_{nn} + \underline{\ell}_{ni}$

J = set of labels of bodies directly connected to body n

$\underline{F}_n$, $\underline{T}_n$ = vector forces and torques externally applied to body n

$\underline{T}_{nj}$ = torques transmitted to body n through hinges connecting n to bodies j

Φ_n^* = augmented inertia dyadic of body n

ω_n = angular velocity vector of body n

The definition of the barycenter vectors, $\underline{b}$, in terms of weighted sums of the $\underline{\ell}$'s, results in the relatively simple form of (1).

A simple rigid body system chosen to approximate the dynamic effects of boom-like appendages on TOPS is shown in Figure 4.

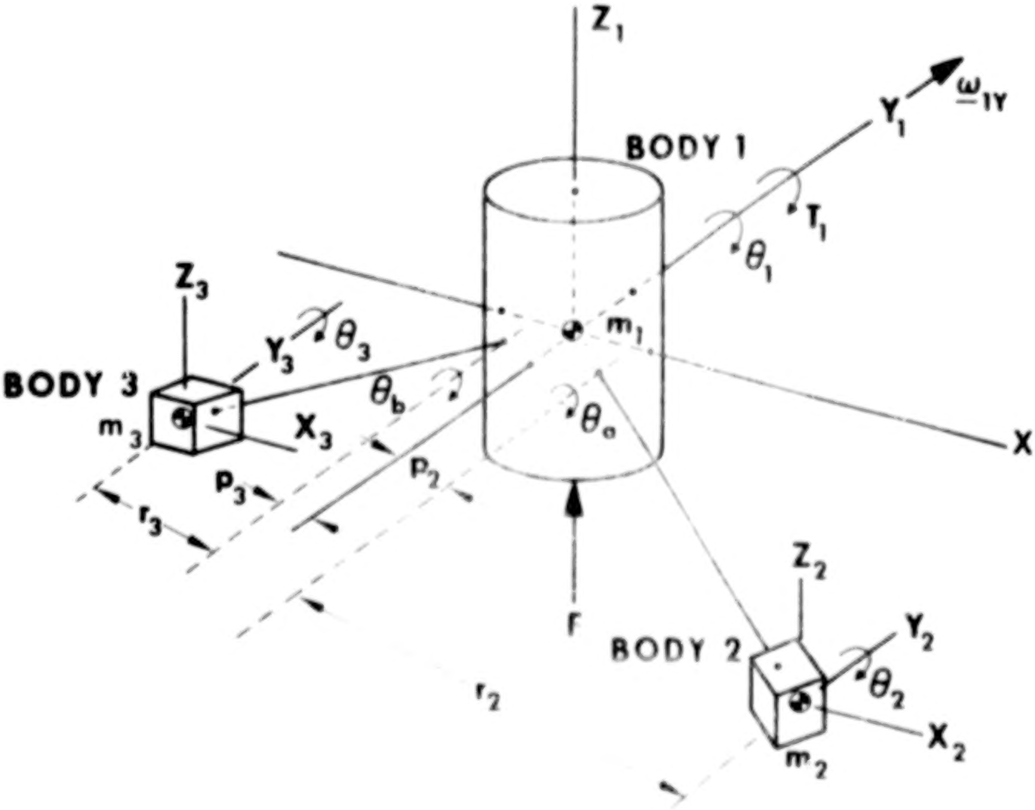

Z_1
$\underline{\omega}_{1Y}$
Y_1
BODY 1
T_1
θ_1
Z_3
θ_3
Y_3
m_1
BODY 3
θ_b
θ_a
m_3
X_3
X_1
P_3
P_2
r_3
Z_2
F
Y_2
BODY 2
θ_2
r_2
X_2
m_2

Figure 4.- Simplified rigid-body model of spacecraft boom dynamics

Applying equation (1) to each of the three rigid bodies results in:

$$\Phi_1^{*} \cdot \dot{\underline{\omega}}_1 + \underline{\omega}_1 \times \Phi_1^{*} \cdot \underline{\omega}_1 = \underline{T}_1 + \underline{T}_{12} + \underline{T}_{13} + \underline{b}_{11} \times \underline{F}_1 \quad (2)$$

$$+ m[\underline{b}_{12} \times (\dot{\underline{\omega}}_2 \times \underline{b}_{21} + \underline{\omega}_2 \times (\underline{\omega}_2 \times \underline{b}_{21}))$$

$$+ \underline{b}_{13} \times (\dot{\underline{\omega}}_3 \times \underline{b}_{31} + \underline{\omega}_3 \times (\underline{\omega}_3 \times \underline{b}_{31}))]$$

$$\Phi_2^{*} \cdot \dot{\underline{\omega}}_2 + \underline{\omega}_2 \times \Phi_2^{*} \cdot \underline{\omega}_2 = \underline{T}_{21} + \underline{b}_{21} \times \underline{F}_1 \quad (3)$$

$$+ m[\underline{b}_{21} \times (\dot{\underline{\omega}}_1 \times \underline{b}_{12} + \underline{\omega}_1 \times (\underline{\omega}_1 \times \underline{b}_{12}))$$

$$+ \underline{b}_{23} \times (\dot{\underline{\omega}}_3 \times \underline{b}_{32} + \underline{\omega}_3 \times (\underline{\omega}_3 \times \underline{b}_{32}))]$$

$$\Phi_3^{*} \cdot \dot{\underline{\omega}}_3 + \underline{\omega}_3 \times \underline{\Phi}_3^{*} = \underline{T}_{31} + \underline{b}_{31} \times \underline{F}_1 \quad (4)$$

$$+ m[\underline{b}_{31} \times (\dot{\underline{\omega}}_1 \times \underline{b}_{13} + \underline{\omega}_1 \times (\underline{\omega}_1 \times \underline{b}_{13}))$$

$$+ \underline{b}_{32} \times (\dot{\underline{\omega}}_2 \times \underline{b}_{23} + \underline{\omega}_2 \times (\underline{\omega}_2 \times \underline{b}_{23}))]$$

where:

$$\underline{T}_2 = \underline{T}_3 = \underline{F}_2 = \underline{F}_3 = 0$$

Assuming small rotations about the hinges (in a direction parallel to Y_1), small $\underline{\omega}$'s (so that second order terms in $\underline{\omega}$ are dropped), approximately collinear body-fixed vector bases in each body, and rotation only in the x-z plane, vector Eqs. (2), (3), and (4) are linearized and reduced to the scalar equations:

$$\Phi_{1yy}\dot{\omega}_{1y} = T_{1y} + T_{12y} + T_{13y} - b_{11x}F_{1z} + m[b_{12x}b_{21x}\dot{\omega}_{2y}$$

$$+ b_{13x}b_{31x}\dot{\omega}_{3y}]$$

$$\phi_{2yy}\dot{\omega}_{2y} = T_{21y} - b_{21x}F_{1z} + m[b_{21x}b_{12x}\dot{\omega}_{1y} + b_{23x}b_{32x}\dot{\omega}_{3y}]$$

$$\phi_{3yy}\dot{\omega}_{3y} = T_{31y} = b_{31x}F_{1z} + m[b_{31x}b_{13x}\dot{\omega}_{1y} + b_{32x}b_{23x}\dot{\omega}_{2y}]$$

By expressing the hinge reaction torques along Y_1 by:

$$T_{12y} = - T_{21y} = k_2(\Theta_2 - \Theta_1) + d_2\ (\dot{\Theta}_2 - \dot{\Theta}_1)$$

$$T_{13y} = - T_{31y} = k_3(\Theta_3 - \Theta_1) + d_3\ (\dot{\Theta}_3 - \dot{\Theta}_1)$$

and Laplace-transforming the scalar differential equations, the transfer functions $\Theta_y(s)/T_{1y}(s)$ may be obtained as:

$$\frac{\Theta_{1y}(s)}{T_{1y}(s)} = \frac{\begin{vmatrix} 1 & (s^2B_{12} - sd_2 - k_2) & (s^2B_{13} - sd_3 - k_3) \\ 0 & (s^2\phi_{2yy} + sd_2 + k_2) & s^2B_{23} \\ 0 & s^2B_{23} & (s^2\phi_{3yy} + sd_3 + k_3) \end{vmatrix}}{s^2\begin{vmatrix} (\phi_{1yy} + B_{12} + B_{13}) & (s^2B_{12} - sd_2 - k_2) & (s^2B_{13} - sd_3 - k_3) \\ (\phi_{2yy} + B_{12} + B_{23}) & (s^2\phi_{2yy} + sd_2 + k_2) & s^2B_{23} \\ (\phi_{3yy} + B_{13} + B_{23}) & s^2B_{23} & (s^2\phi_{3yy} + sd_3 + k_3) \end{vmatrix}}$$

or

$$\frac{\theta_{1y}(s)}{T_{1y}(s)} = \frac{N_4s^4 + N_3s^3 + N_2s^2 + N_1s + N_o}{s^2(D_4s^4 + D_3s^3 + D_2s^2 + D_1s + D_o)}$$

where:

$$B_{12} = - mb_{12x}b_{21x} \qquad \phi_{1yy} = I_{1yy} + m_1b_{11x}^2 + m_2b_{12x}^2 + m_3b_{13x}^2$$

$$B_{13} = - mb_{13x}b_{31x} \qquad \phi_{2yy} = I_{2yy} + m_1b_{21x}^2 + m_2b_{22x}^2 + m_3b_{23x}^2$$

$$B_{23} = - mb_{23x}b_{32x} \qquad \phi_{3yy} = I_{3yy} + m_1b_{31x}^2 + m_2b_{32x}^2 + m_3b_{33x}^2$$

It mav be verified that the coefficient values for numerator and denominator polynominals are as follows:

$$N_4 = \phi_{2yy}\phi_{3yy} - B_{23}^2$$

$$N_3 = \phi_{2yy}d_3 + \phi_{3yy}d_2$$

$$N_2 = \phi_{2yy}k_3 + \phi_{3yy}k_2 + d_2d_3$$

$$N_1 = d_2k_3 + d_3k_2$$

$$N_0 = k_2k_3$$

$$D_4 = \phi_{1yy}\phi_{2yy}\phi_{3yy} - B_{12}^2\phi_{3yy} - B_{13}^2\phi_{2yy} + 2B_{12}B_{13}B_{23}$$

$$\begin{aligned} D_3 = {} & (\phi_{1yy}\phi_{2yy} + \phi_{2yy}\phi_{3yy} + 2B_{13}\phi_{2yy} - 2B_{12}B_{23} - B_{12}^2 - B_{23}^2)d_3 \\ & + (\phi_{1yy}\phi_{3yy} + \phi_{2yy}\phi_{3yy} + 2B_{12}\phi_{3yy} - 2B_{13}B_{23} - B_{13}^2 - B_{23}^2)d_2 \end{aligned}$$

$$\begin{aligned} D_2 = {} & (\phi_{1yy}\phi_{2yy} + \phi_{2yy}\phi_{3yy} + 2B_{13}\phi_{2yy} - 2B_{12}B_{23} - B_{12}^2 - B_{23}^2)k_3 \\ & + (\phi_{1yy}\phi_{3yy} + \phi_{2yy}\phi_{3yy} + 2B_{12}\phi_{3yy} - 2B_{13}B_{23} - B_{13}^2 - B_{23}^2)k_2 \\ & + (\phi_{1yy} + \phi_{2yy} + \phi_{3yy} + 2B_{12} + 2B_{13} + 2B_{23})d_2d_3 \end{aligned}$$

$$D_1 = (\phi_{1yy} + \phi_{2yy} + \phi_{3yy} + 2B_{12} + 2B_{13} + 2B_{23})(d_2k_3 + d_3k_2)$$

$$D_0 = (\phi_{1yy} + \phi_{2yy} + \phi_{3yy} + 2B_{12} + 2B_{13} + 2B_{23})k_2k_3$$

It is apparent that if the term

$$I_T = \phi_{1yy} + \phi_{2yy} + \phi_{3yy} + 2B_{12} + 2B_{13} + 2B_{23}$$

is factored out of the denominator, the dc or Bode gain of the ratio of polynomials is unity. The value, I_T, happens to be the y-axis moment of inertia of the entire system about its center of mass.

An initial look at the science instrumentation and RTG boom structures provided the following numerical result, where rotations are now assumed about the x_1 axis only and the relations developed above are used with appropriate substitutions:

$m_1 = 25.6$ slugs, $m_2 = 8.72$ slugs, $m_3 = 5.4$ slugs

$m = 39.72$ slugs $\qquad \rho_2 = 1.97$ ft

$I_1 = 235.6$ slug-ft^2 $\qquad \rho_3 = 1.92$ ft

$I_2 = 62.4$ slug-ft^2 $\qquad r_2 = 7.53$ ft

$I_3 = 61.2$ slug-ft^2 $\qquad r_3 = 6.89$ ft

Thus:

$\phi_{1xx} = 288.180 \qquad B_{12} = 118.095$

$\phi_{2xx} = 448.286 \qquad B_{13} = 77.815$

$\phi_{3xx} = 282.698 \qquad B_{23} = 61.506$

Given also are the natural frequencies of vibration and damping for bodies 2 and 3 about their respective connecting hinges (assuming body 1 is fixed in inertial space):

$\omega_2 = 2\,\pi\,(4.2)$ rad/sec $\qquad \delta_2 = .005$

$\omega_3 = 2\,\pi\,(7.1)$ rad/sec $\qquad \delta_3 = .005$

so that

$$I_{E2}\ddot{\Theta}_2 + d_2\dot{\Theta}_2 + k_2 = 0$$

$$I_{E3}\ddot{\Theta}_3 + d_3\dot{\Theta}_3 + k_3 = 0$$

The values of k_2, k_3, d_2, and d_3 are then computed from the relations above:

$k_2 = 3.8778\ (10^5)$ ft-lb/rad, $d_2 = 146.944$ ft-lb/rad/sec

$k_3 = 6.3196\ (10^5)$ ft-lb/rad, $d_3 = 141.661$ ft-lb/rad/sec

As a result, the polynomial coefficients can be computed to give:

$I_T = 1534.$ slug-ft^2

$N_4' = 1.2294657$ $D_4' = .20204871$

$N_3' = 1.0504542$ $D_3' = .51577439$

$N_2' = 3929.4427$ $D_2' = 1855.0591$

$N_1' = 1465.0603$ $D_1' = 1465.0603$

$N_0' = 2.450614\ (10^6)$ $D_0' = 2.450614\ (10^6)$

Pole and zero locations of an autopilot control loop incorporating the structural transfer function just derived are given in Figure 5. Included are pairs of complex poles representing gyro dynamics and gimbaled motor actuator dynamics, a real zero from the gyro response, and a real filter pole. The autopilot presented here is taken directly from the Mariner '71 design. A computer program was used to generate a root locus diagram of the system, and the result is given in Figure 6. The accompanying gain locus of Figure 7 indicates a Bode gain of about 6.8 at the crossover point of the lower locus into the right-half plane. Meanwhile the lowest structural pole (at $-.265 \pm j40$) does not enter the instability region until a Bode gain of about 48 has been reached.

10 OPEN LOOP POLES

	REAL	IMAGINARY	ORDER	
AUTOPILOT →	−9.0000E 00	0.0000E−399	1	
	−8.8000E 00	1.4200E 01	1	GIMBAL SERVO
	−8.8000E 00	−1.4200E 01	1	
GYRO	−3.0800E 01	−8.2400E 01	1	
	−3.0800E 01	8.2400E 01	1	
	0.0000E−39	0.0000E −39	2	← RIGID BODY
	−1.0116E 00	−8.7058E 01	1	STRUCTURE (RTG AND SCIENCE BOOMS)
	−1.0116E 00	8.7058E 01	1	
	−2.6475E−01	−4.0000E 01	1	
	−2.6475E−01	4.0000E 01	1	

5 OPEN LOOP ZERO(S)

	REAL	IMAGINARY	ORDER	
GYRO →	−5.6500E−01	0.0000E−39	1	
	−2.7161E−01	−4.8439E 01	1	STRUCTURE
	−2.7161E−01	4.8439E 01	1	
	−1.5559E−01	−2.9146E 01	1	
	−1.5559E−01	2.9146E 01	1	

Figure 5.- Gimbaled autopilot open loop root locations

In the same manner, one can look at the autopilot system about the y-axis, where magnetometer boom bending enters quite significantly. In this case, the large science and RTG structures are assumed to be virtually rigid (torsionally stiff) an simply become a part of the rigid central body.

Structural data for this configuration are as follows:

m_2 = .183 slugs

m_3 = .183 slugs

I_1 = 193.7 slug-ft^2

I_2 = 17. slug-ft^2

I_3 = 17. slug-ft^2

m_1 = 39.35 slugs

m = 39.72 slugs

ρ_2 = 1.1 ft

ρ_3 = 1.1 ft

r_2 = 19.4 ft

r_3 = 19.4 ft

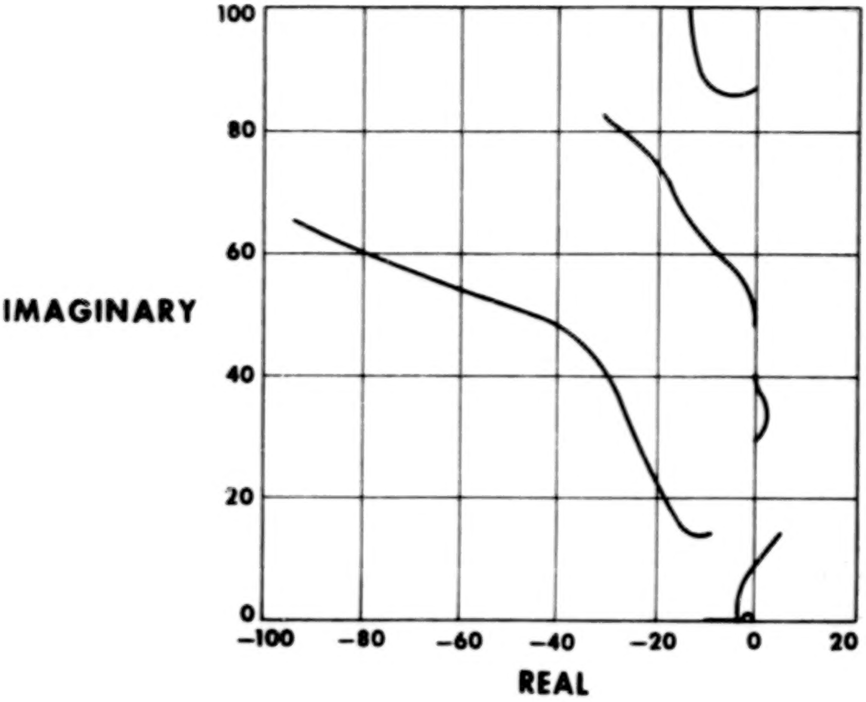


Figure 6.- Autopilot loop root locus
with science and RTG booms

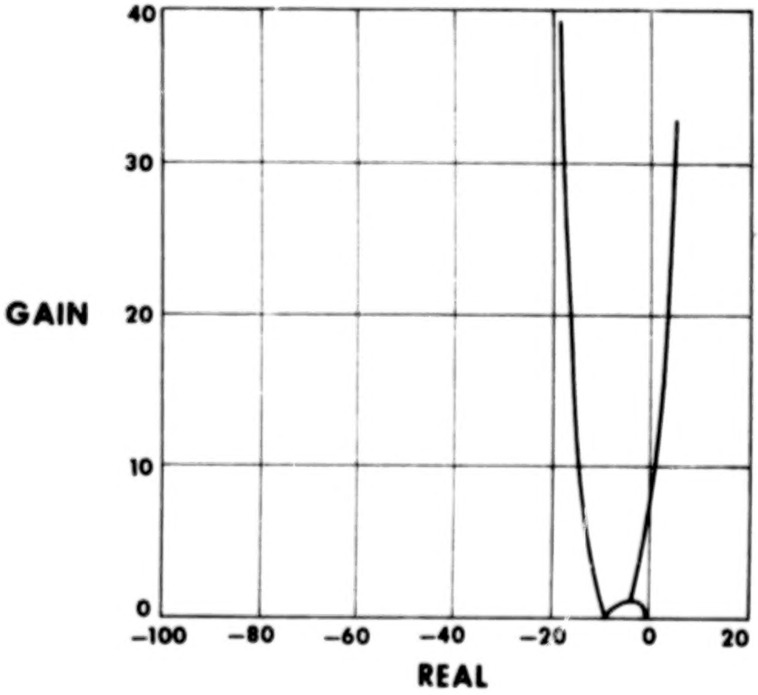


Figure 7.- Autopilot loop gain locus
(science and RTG booms)

$\omega_2 = 2\pi\,(.52)$ rad/sec $\qquad \delta_2 = .005$

$\omega_3 = 2\pi\,(.57)$ rad/sec $\qquad \delta_3 = .005$

As a result, the corresponding boom stiffness and damping factors may be computed as before:

$k_2 = 1002.38$ ft-lb/rad $\qquad d_2 = 3.06795$ ft-lb/rad/sec

$k_3 = 1204.41$ ft-lb/rad $\qquad d_3 = 3.36295$ ft-lb/rad/sec

Additional constraints required are:

$\phi_{1yy} = 194.14286$ slug-ft^2 , $B_{12} = 3.90523$ slug-ft^2

$\phi_{2yy} = 85.5564$ slug-ft^2 , $B_{13} = 3.90523$ slug-ft^2

$\phi_{3yy} = 85.5564$ slug-ft^2 , $B_{23} = .31732$ slug-ft^2

Factoring out the new $I_T = 381.5$ slug-ft^2 from the denominator and factoring 10^3 from both numerator and denominator results in:

$N_4' = 7.319797 \qquad D_4' = 3.718233$

$N_3' = .5502047 \qquad D_3' = .4143499$

$N_2' = 188.8153 \qquad D_2' = 142.1963$

$N_1' = 7.0660235 \qquad D_1' = 7.0660235$

$N_0' = 1207.2765 \qquad D_0' = 1207.2765$

A root locus diagram of the autopilot discussed above and incorporating the new y-axis dynamics is shown in Figure 8. The loci for this case are significantly different in character due to the nearness of the structural poles and zeros to the origin. The structural pole now crosses into the right-half plane at a gain of about 2.6 (see Figure 9).

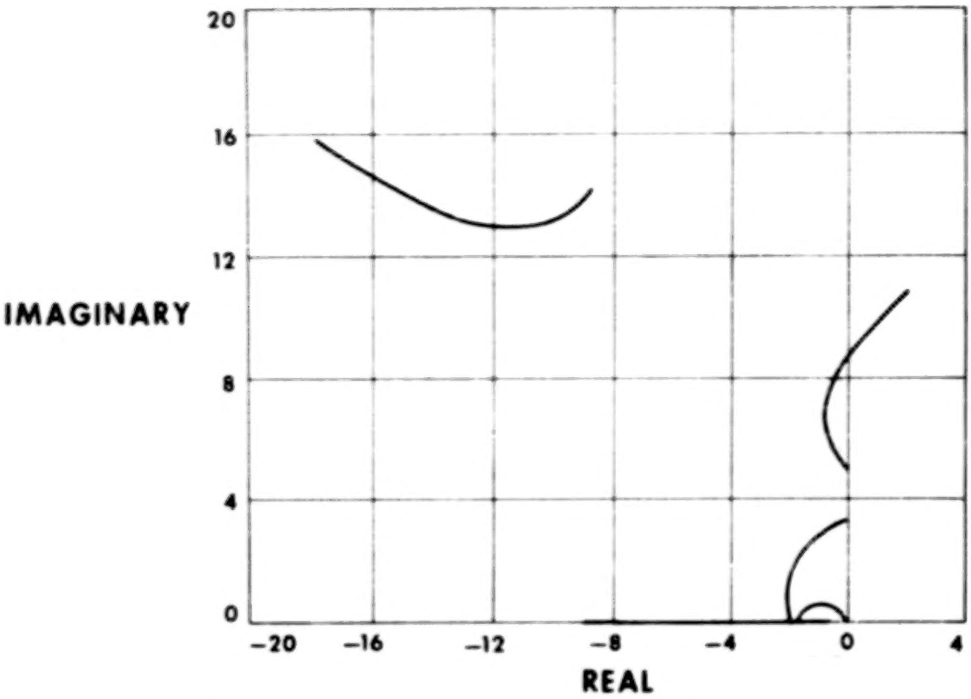


Figure 8.- Autopilot root locus
with magnetometer booms

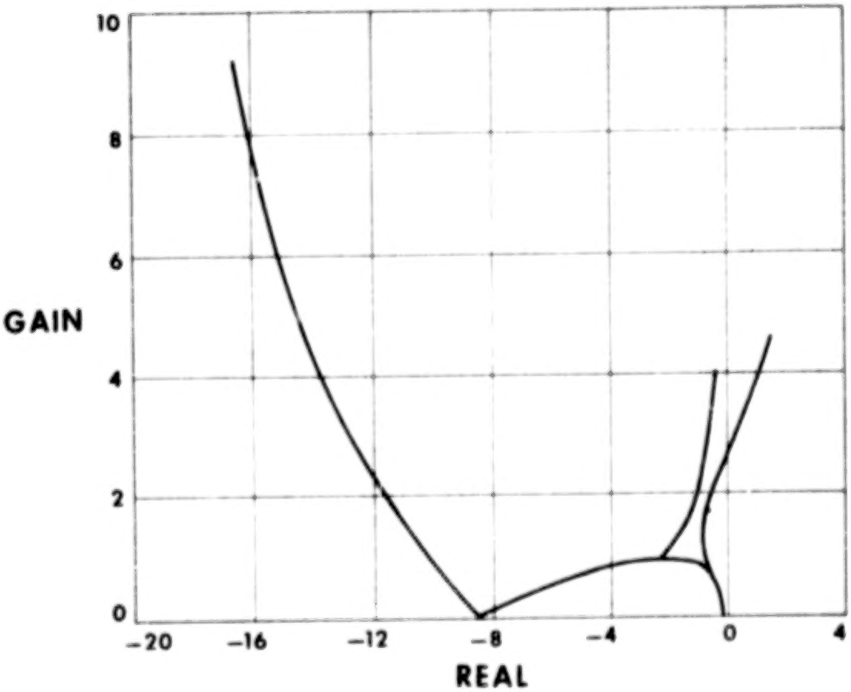


Figure 9.- Autopilot gain locus
(magnetometer booms)

Finally, one can compare the autopilot loop for the rigid body (neglecting appendage vibrations) to the results above. Figure 10 is a root locus of the system minus the effect of bending members on the spacecraft. Gain crossover occurs at 6.3 (see Figure 11) with the main locus exhibiting almost the same appearance as in the case for relatively large but stiff appendages. Apparently then, from this crude first look at the effect of flexible members, the autopilot's critical gain is reduced by a factor of about 2.5 when the influence of the low-frequency magnetometer booms is included, but not significantly affected by the relatively stiff RTG and science instrumentation appendages. The major influence is felt by yaw since its total inertia is small compared to pitch (or roll) and therefore has a greater percentage of that inertia contributed by flexible members.

HYBRID COORDINATE MODEL

A great deal of experience and capability has been developed in the analysis of flexible structures using modal deformation coordinates, and they have long been efficiently applied to aircraft, missiles, and spacecraft as well. From the control standpoint, however, a strictly modal coordinate approach is not capable of accommodating rotors, nonlinear internal controls, discrete dampers, etc. At the same time, the opposite approach of modeling the system as a large collection of rigid bodies appears to be extremely cumbersome. Computer solutions to the resulting equations may be forced to include irrelevant high-frequency components, drastically increasing run times and therefore cost. Typically, the solution of large matrix equations is also required at each numerical integration step, aggravating even further the computational problem.

Recently, Likins (refs. 3, 4, 5) has developed in detail a "hybrid coordinate" approach to the vehicle control problem when flexible appendages are involved. It is based on the assumption that flexible portions of the craft may be accurately characterized by small, linearly elastic deformations and are therefore subject to modal analysis. Furthermore, if the flexible appendages are attached to a rigid base which undergoes small motions or which has nearly constant angular velocity, normal modes of vibration can be shown to exist. Discrete coordinates are retained for describing the rigid base motion or the motion of other essentially rigid components, such as gimbaled engines, rotors, discrete dampers, and the like.

For example, if the spacecraft's rigid bus is assumed to experience a sufficiently small inertial angular velocity, so

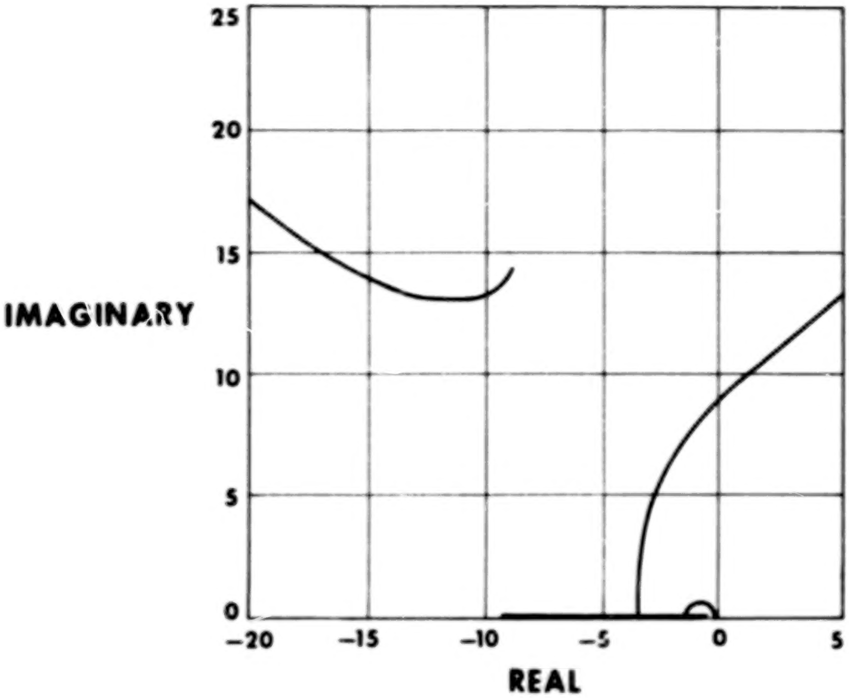


Figure 10.- Autopilot loop root locus for entirely rigid spacecraft

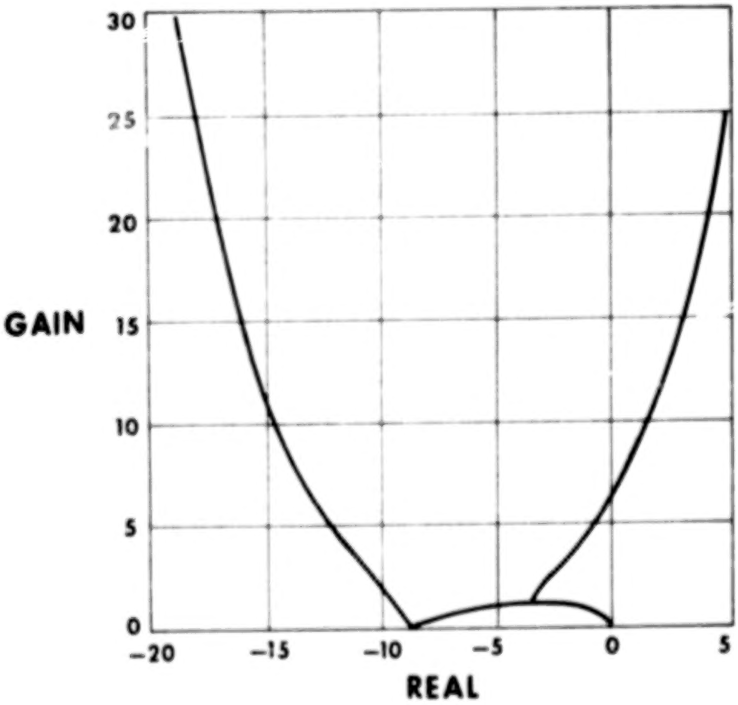


Figure 11.- Gain locus for rigid-body autopilot loop

that second-degree terms and higher derivatives in this variable are negligible, the completely linearized matrix equation describing the flexible appendage becomes

$$M[E - \Sigma_E \Sigma_E^T M/m]\,\ddot{q} + Kq = -M\,\Sigma_E[\ddot{X} + \ddot{e} - 2\dot{e}\tilde{\dot{\Theta}} - (\tilde{e} + \tilde{R})\ddot{\Theta}] + M\tilde{r}\Sigma_E\ddot{\Theta} + \lambda \tag{5}$$

where:

M = mass matrix of the flexible appendage (3nx3n) (n is the number of point masses used to model the appendage)
m = total vehicle mass (scalar)
q = column matrix deformation coordinates (3nx1)
K = appendage stiffness matrix (3nx3n)
X = inertial position of the system center of mass (3x1)
e = position of system mass center with respect to a point fixed in rigid base (bus), not including flexible appendage contributions (3x1)
R = location of appendage attachment point in rigid base (3x1)
$\tilde{r}, \tilde{R}$ = skew symmetric matrices locating masses in the flexible appendage (3nx3n, 3x3)
Θ = inertial attitude of rigid base body (3x1)
E = identity matrix, any order
Σ_E = matrix $[EEE...E]^T$, (3nx3)
λ = matrix of external forces applied to appendage (3nx1).

If in Eq. (5) the transformation $q = \Phi\eta$ is substituted followed by a premultiplication of the equation by Φ^T, where Φ is the matrix of eigenvectors of the homogeneous equation, the coefficient matrices of $\eta, \ddot{\eta}$ are diagonalized. Suitable normalization of the eigenvectors and insertion of a diagonal structural damping matrix results in

$$\ddot{\eta} + 2\,\xi\omega\dot{\eta} + \omega^2\eta = -\Phi^T M\,\{[-\Sigma_E(\tilde{e} + \tilde{R}) - \tilde{r}\,\Sigma_E]\ddot{\Theta} + \Sigma_E(\ddot{X} + \ddot{e} - 2\,\dot{e}\tilde{\dot{\Theta}})\} + \Phi^T\lambda \tag{6}$$

Equation 6 may then be reduced in dimension (η's are uncoupled) to eliminate those modal coordinates η with unwanted high frequencies.

In addition to appendage Eq. (6), a vector equation of motion can be written for the entire vehicle from the basic relation $\underline{T} = \dot{\underline{H}}$, which in matrix form becomes

$$[I - I_w E]\ddot{\theta} + (\hat{R}\Sigma_E^T + \Sigma_E^T r) M \Phi \ddot{\eta} = T - \tau - \tilde{\dot{\theta}} I_w \dot{\phi} \qquad (7)$$

where:

I = total undeformed vehicle inertia matrix (3x3)

I_w = spin-axis moment of inertia for each of 3 identical momentum wheels whose spin axes are parallel to the reference axes (scalar)

τ = $[\tau_1 \tau_2 \tau_3]^T$, momentum wheel applied torques

$\dot{\phi}$ = $[\dot{\phi}_1 \dot{\phi}_2 \dot{\phi}_3]^T$, momentum wheel spin rates

T = matrix or torques applied to rigid base (3x1).

For the case of the TOPS autopilot, e and its derivatives were considered small, external forces (λ) were not applied to the flexible appendages, and the momentum wheel system was assumed temporarily inoperative. Equations (6) and (7) could then be written in combined matrix form as:

$$\left[\begin{array}{c|c} I & \bar{\Delta}^T \\ \hline \bar{\Delta} & E \end{array}\right] \left\{\begin{array}{c} \ddot{\theta} \\ \hline \ddot{\bar{\eta}} \end{array}\right\} = \left\{\begin{array}{c} T \\ \hline - \bar{\Phi}^T M \Sigma_E \frac{F}{m} = 2 \xi \omega \dot{\bar{\eta}} - \omega^2 \bar{\eta} \end{array}\right\} \qquad (8)$$

where the relation $\ddot{X} = F/m$ is used; F_T = matrix of force components applied to rigid vehicle, 3x1; $\bar{\Delta} = - \bar{\Phi}^T M(\Sigma_E R + r\Sigma_E)$, N x 3; and the over bar indicates the result of modal truncation.

The implementation of Eq. (8) for a computer simulation program requires first obtaining the torque vector as a function of the θ matrix vector and in terms of the vehicle geometry. This is rather straight-forward, requiring a reproduction of essentially the feedback branch of Figure 2 for both pitch and yaw and the addition of a simple bang-bang controller with derived-rate feedback for roll. Torque components are formed considering the gimbal mounting geometry, misalignments, and center of mass location. Location of the vehicle's center of mass will of course be affected by flexible appendage deformations and these were included in the program by the addition of the term: $-\Sigma_E^T \frac{\underline{M}}{m} \bar{\Phi} \bar{\eta}$

to the nominal c.m. location of the undeformed vehicle. Reaction torques on the rigid portion of the spacecraft due to the gimbaled engine were not taken into account, i.e., the engine was assumed to be massless.

It remains then to obtain from the structural analyst specific values for the Δ and $\Phi^T M \Sigma_E$ matrices, to appropriately truncate these after an examination of the modal frequencies, and to incorporate the result in the computer program. For example, Table I lists part of the structural data obtained from a hybrid coordinate model of the TOPS configuration shown in Figure 1. The discrete mass model (point masses) also provides the total undeformed vehicle's inertia matrix which is given below:

$$I = \begin{bmatrix} 1139. & -1.694 & -2.041 \\ -1.694 & 347.9 & -25.75 \\ -2.041 & -25.75 & 1264. \end{bmatrix}, \text{ slug-ft}^2$$

Also, $m = 38.54$ slugs.

RESULTS OF THE THREE-AXIS SIMULATIONS

Truncation of the modal information given in Table I resulted in the retention of modes 1 through 5, including the effects of an antenna torsion mode of vibration characterized by the frequency at 1.16 Hz. On the basis of the root locus information for the relatively high frequency RTG and science instrumentation booms, some justification can be made to ignore modal frequencies above 5 Hz since they showed little effect on autopilot stability. The mode at 3.85 Hz is somewhat borderline until simulation run times with and without the 6th mode are compared. Computer integration speed is slowed by a factor of about 5 if this mode is retained, and it was not considered a significant loss in model accuracy to retain only 5 modes.

Control nonlinearities which were included as part of the detailed, three-axis computer simulation are (1) the controlled saturation built into the autopilot compensation (Figure 3); (2) gimbal, servo-drive amplifier, voltage saturation; and (3) the roll-axis, bang-bang control system.

Figures 12 through 27 illustrate the kind of vehicle responses from which striking comparisons can be made between a rigid idealization of the TOPS vehicle and a detailed structural model of its flexibility. Four variables are plotted

TABLE I.- TOPS HYBRID-COORDINATE STRUCTURAL MODEL DATA

Mode	Freq. (Hz)	$\bar{\Phi}^T M \Sigma_E$ (slug)			$\bar{\Lambda}$ (slug-ft)		
		X	Y	Z	X	Y	Z
1	0.74	.6381E-4	.7994E-4	.4092E00	-.6989E01	-.1150E02	.3856E-2
2	0.75	.3130E00	.3519E00	-.1537E-3	-.3933E00	.3582E00	.1329E02
3	0.76	.2964E00	-.3620E00	-.1059E-4	.4072E00	.3357E00	.1254E02
4	0.76	-.3970E-4	-.4834E-4	.5278E00	-.8966E01	.8642E01	-.1440E-3
5	1.16	.2777E-1	.1248E-2	.1468E-2	-.2132E-1	.1469E-1	.4602E01
6	3.85	-.2733E01	-.30035E-1	-.1554E-2	.1673E00	.6115E00	-.2636E02
7	5.02	.4006E00	.4416E-1	.2425E-1	-.6592E00	.4022E01	.1107E01
8	5.66	-.4561E-2	-.2416E-2	-.1875E00	.8579E00	.1944E01	-.7297E-1
9	5.66	.2074E00	.1776E00	-.7461E-2	-.1245E00	.2310E00	.2673E01
10	5.69	-.1780E-1	.9784E-2	.2416E00	-.1100E01	.1477E01	-.2134E00

which give some idea of the vehicle response under the conditions of: (1) completely rigid spacecraft with autopilot loop Bode gain of 7 dB (Figures 12 through 15), (2) flexible craft also with a 7 dB-loop gain in both pitch and yaw axes (Figures 16 through 19), (3) rigid craft with loop gain lowered to 0 dB (Figures 20 through 23), and (4) flexible craft with 0-dB gain (Figures 24 through 27). In all cases, the transient disturbance is due to a c.m. offset in the x direction of .02 ft.

The flexible spacecraft with 7-dB loop gain in pitch and yaw is seen to be unstable (yaw-gimbal, servo-drive amplifier is saturated, limiting the output gimbal angle), although the rigid body root loci and Bode design efforts resulted in choosing 7 dB as the nominal gain design point. Actually, the pitch loop would remain stable at 7 dB provided the yaw gain was reduced. A gain of 1 (0 dB) seems to provide satisfactory response, although the stability margin is low.

CONCLUSIONS

A practical application of the hybrid coordinate approach to flexible vehicle dynamic modeling has been demonstrated with respect to the TOPS configuration. Thus far its value has been proven as a reasonably precise tool for testing flexible body control system designs. Although a moderate amount of detailed work must be carried out by structural analysts, computer implementation of the hybrid coordinate system of equations is simple and straightforward. The attitude control analyst is provided with explicit access to the rigid central body which is generally the object of his efforts. As a result, any type of controller may be attached to this body so long as the rigid body's motion may be reasonably considered small.

Not unexpectedly, results of the detailed 3-axis simulations show significantly less stability margin even though certain stabilizing steps were taken, i.e., a much more highly damped gimbal actuator was used in addition to somewhat more stiff magnetometer booms (.75 Hz as opposed to .52 Hz in the simple flexible model). This appears to be the result of interaxis coupling effects which could be verified by eigenvalue analyses.

As yet the hybrid coordinate structural model has not been used to predict antenna pointing errors, science pointing errors, or other quantities based on actual structural deflections, although this could easily be done.

Likins, in references 4 and 5, has shown that the modal data produced by the hybrid coordinate approach is well suited to root locus studies, and present work at JPL is centered on

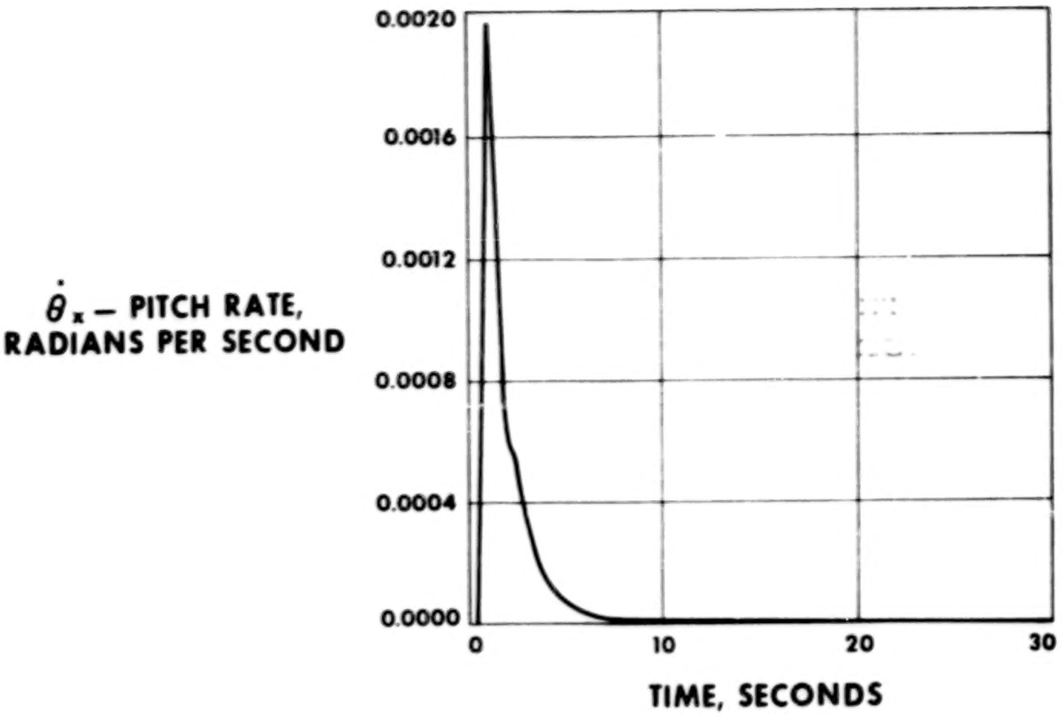


Figure 12.- TOPS, pitch rate vs time, rigid vehicle (gain = 7 dB)

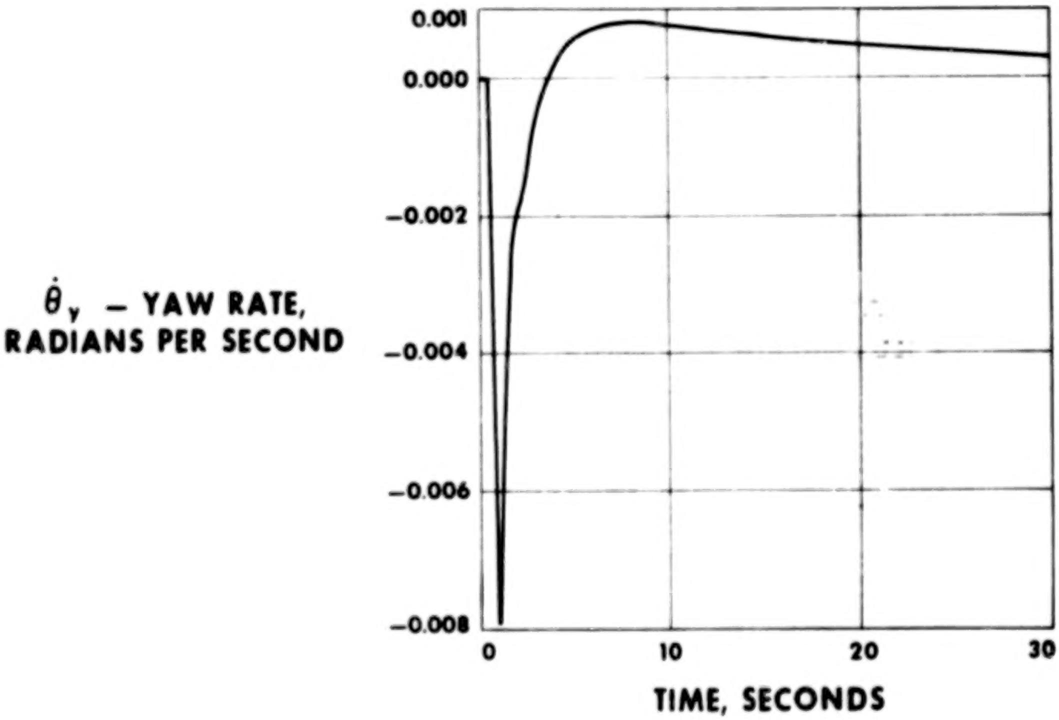


Figure 13.- Yaw rate vs time, rigid vehicle (gain = 7 dB)

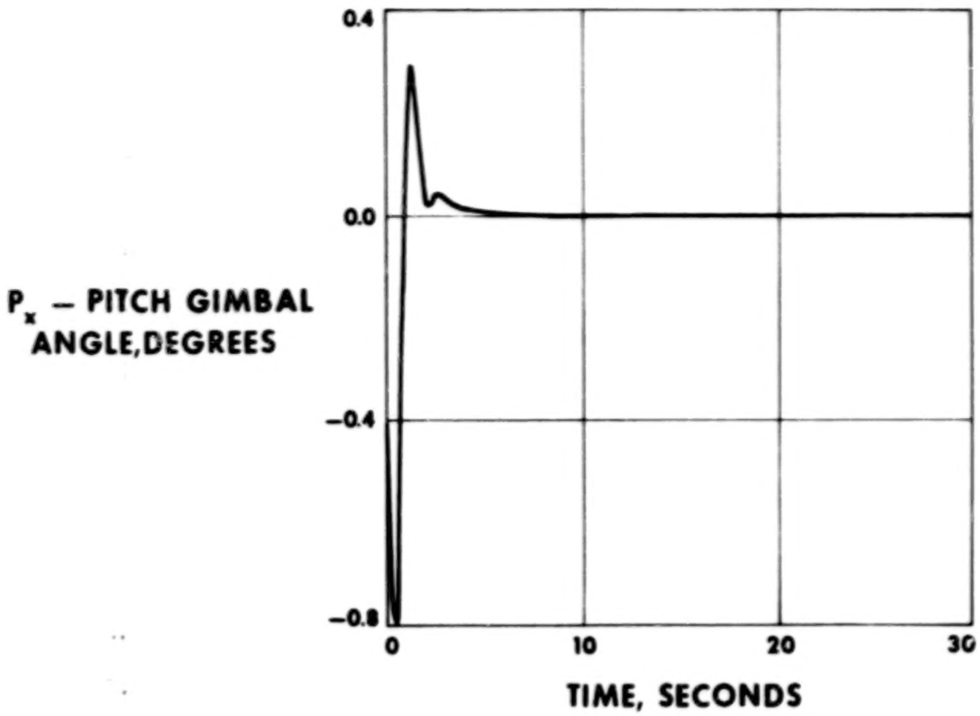


Figure 14.- Pitch gimbal angle vs time, rigid vehicle (gain = 7 dB)

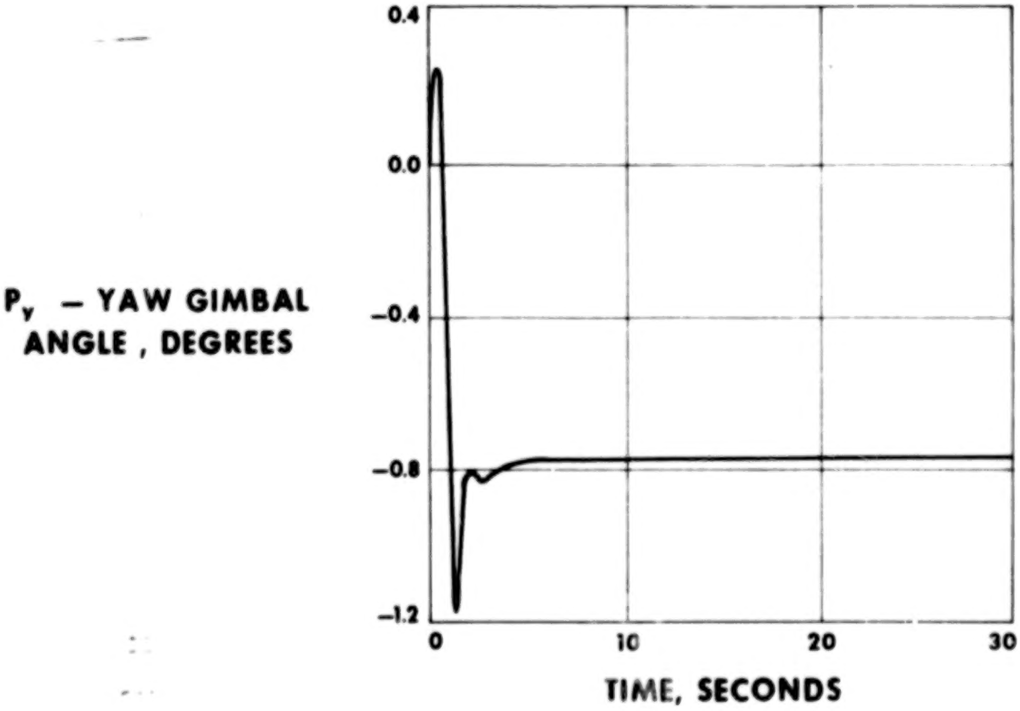


Figure 15.- Yaw gimbal angle vs time, rigid vehicle (gain = 7 dB)

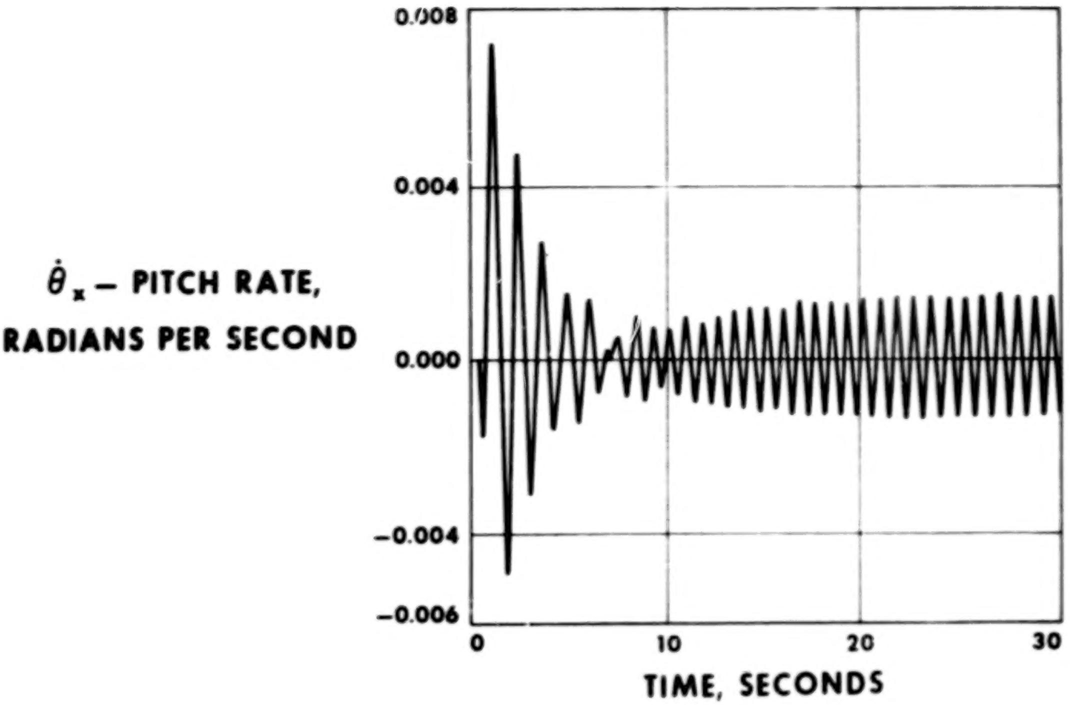


Figure 16.- Pitch rate vs time, flexible vehicle
(gain = 7 dB)

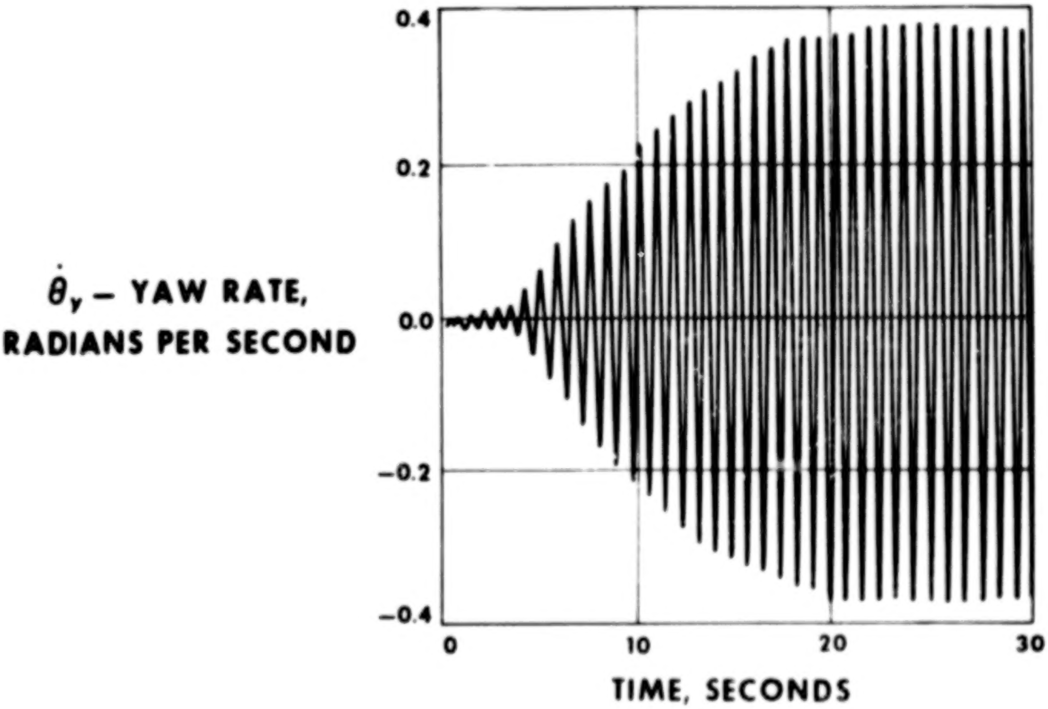


Figure 17.- Yaw rate vs time, flexible vehicle
(gain = 7 dB)

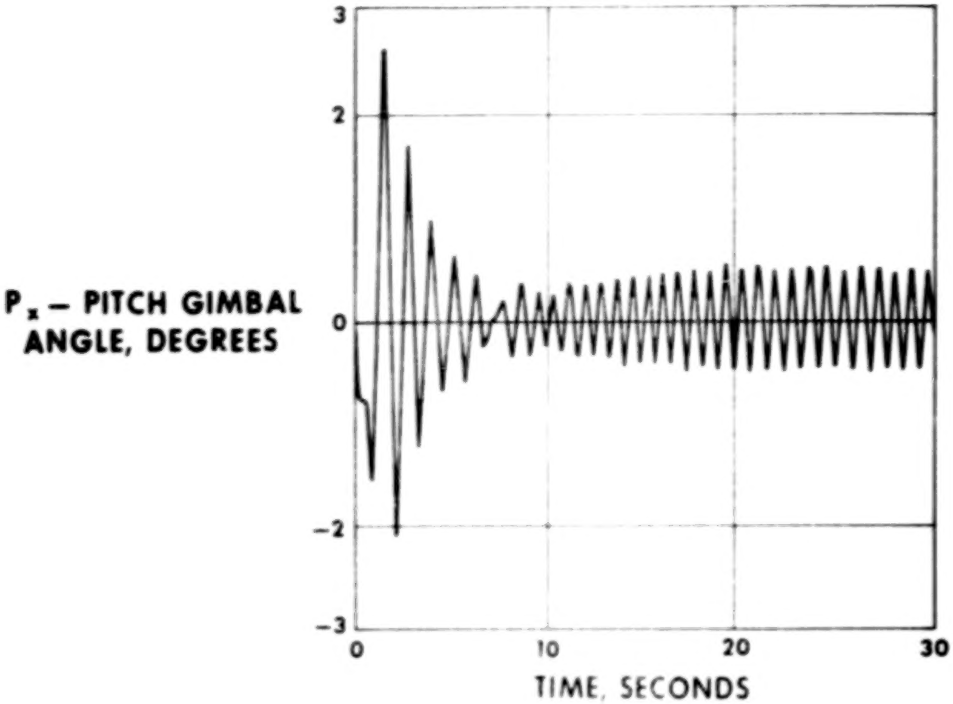


Figure 18.- Pitch gimbal angle vs time, flexible vehicle (gain = 7 dB)

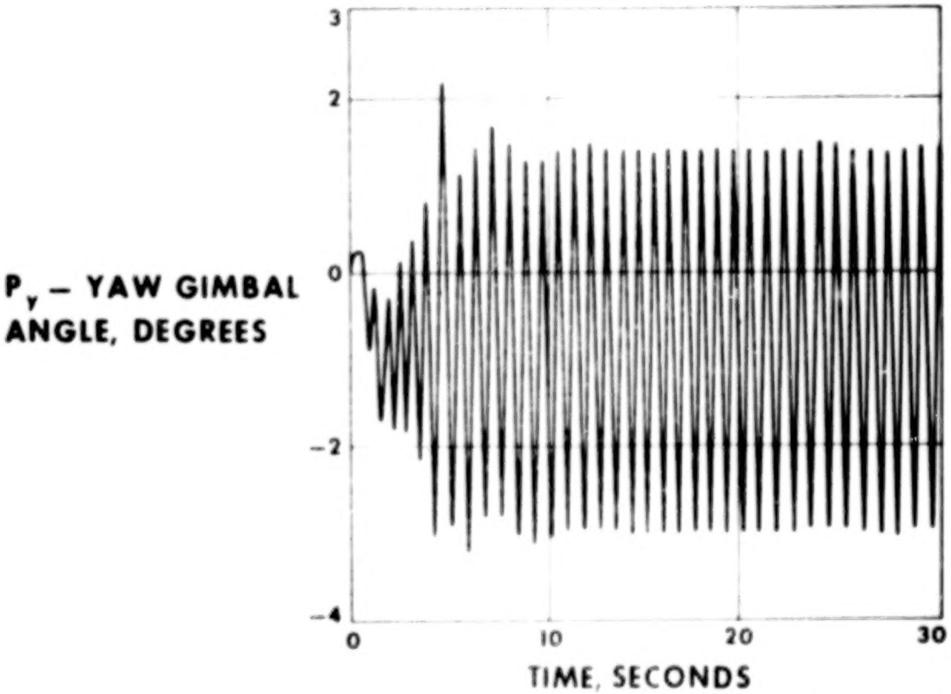


Figure 19.- Yaw gimbal angle vs time, flexible vehicle (gain = 7 dB)

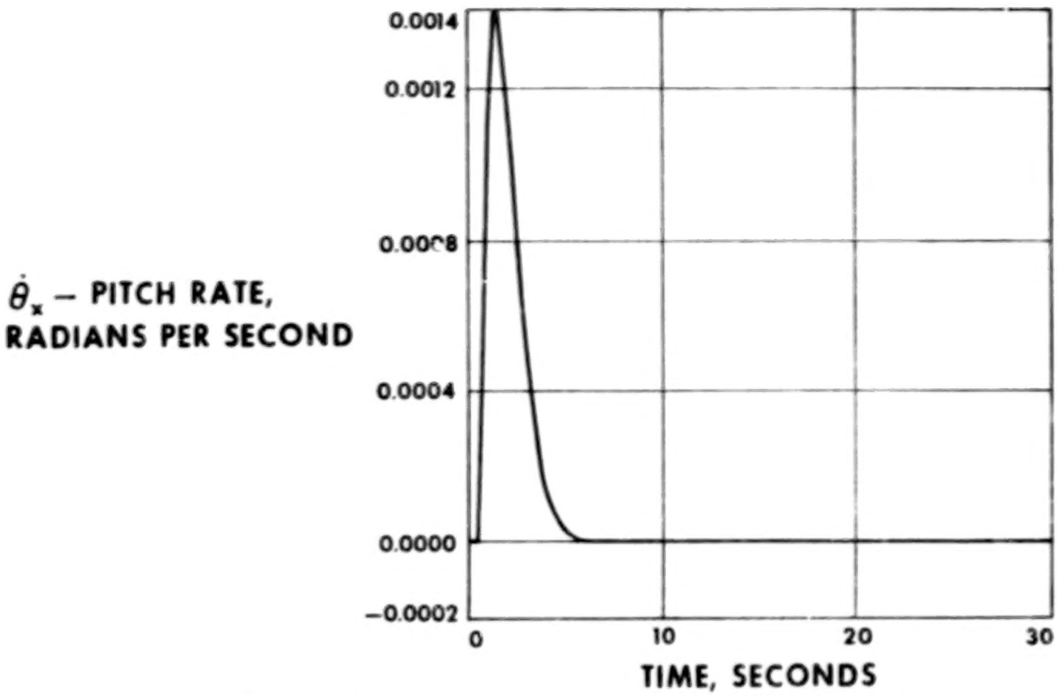


Figure 20.- Pitch rate vs time, rigid vehicle
(gain = 0 dB)

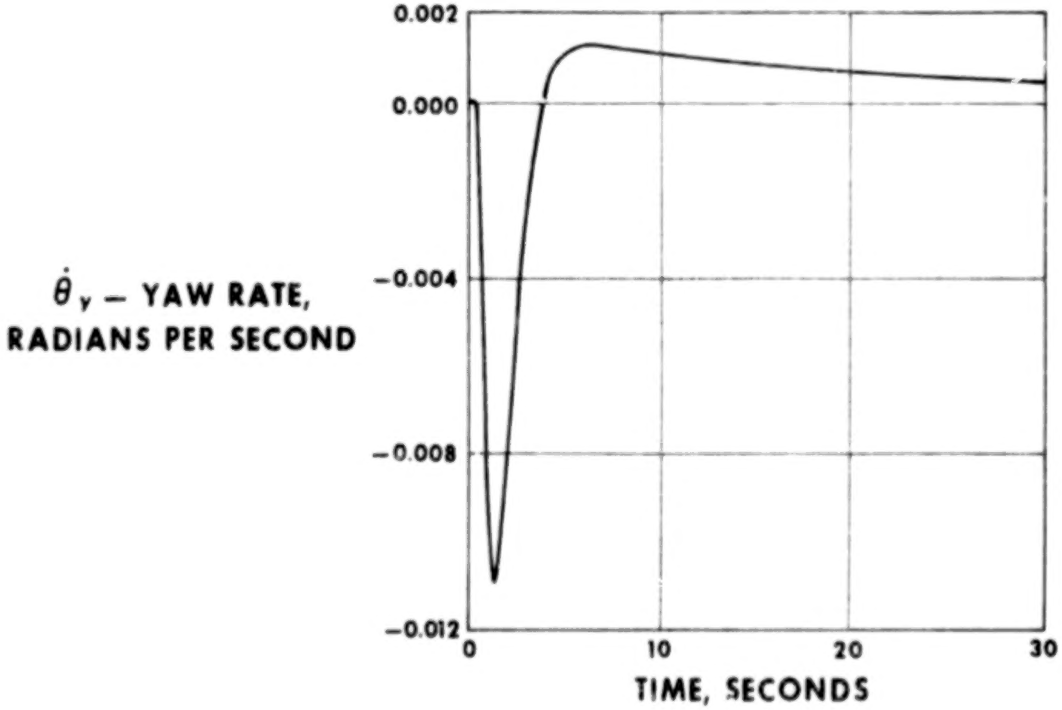


Figure 21.- Yaw rate vs time, rigid vehicle
(gain = 0 dB)

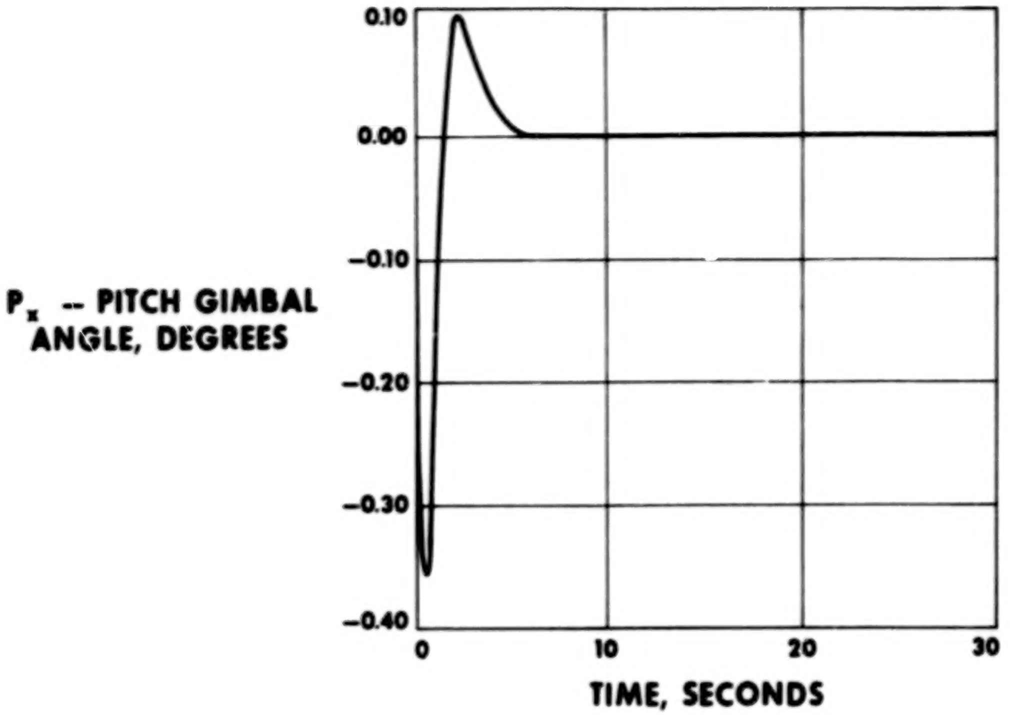


Figure 22.- Pitch gimbal angle vs time, rigid vehicle (gain = 0 dB)

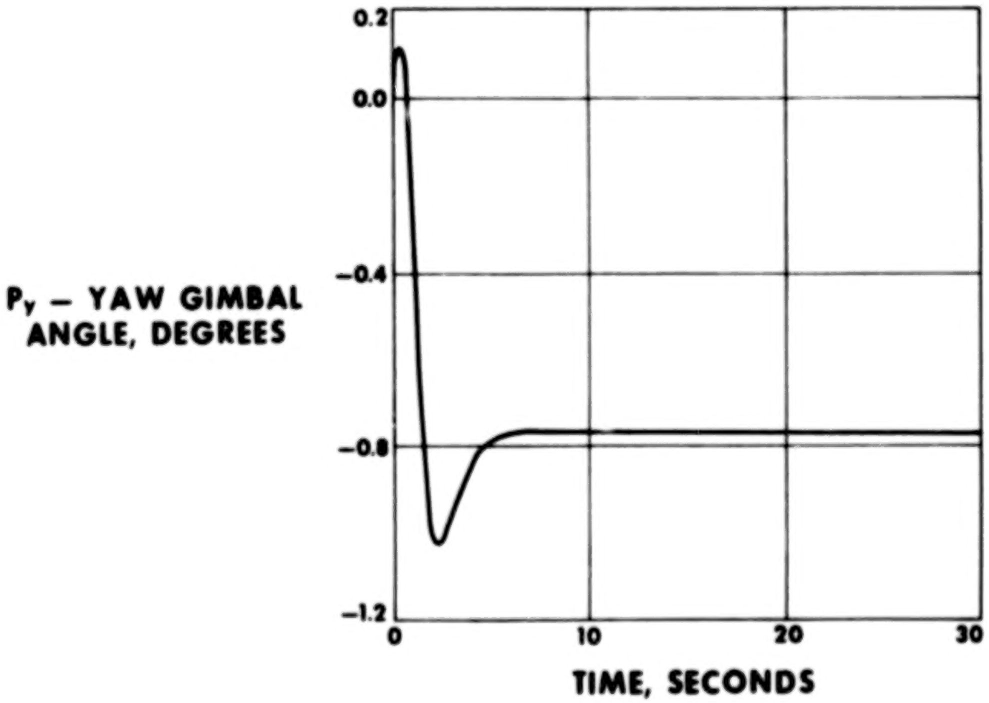


Figure 23.- Yaw gimbal angle vs time, rigid vehicle (gain = 0 dB)

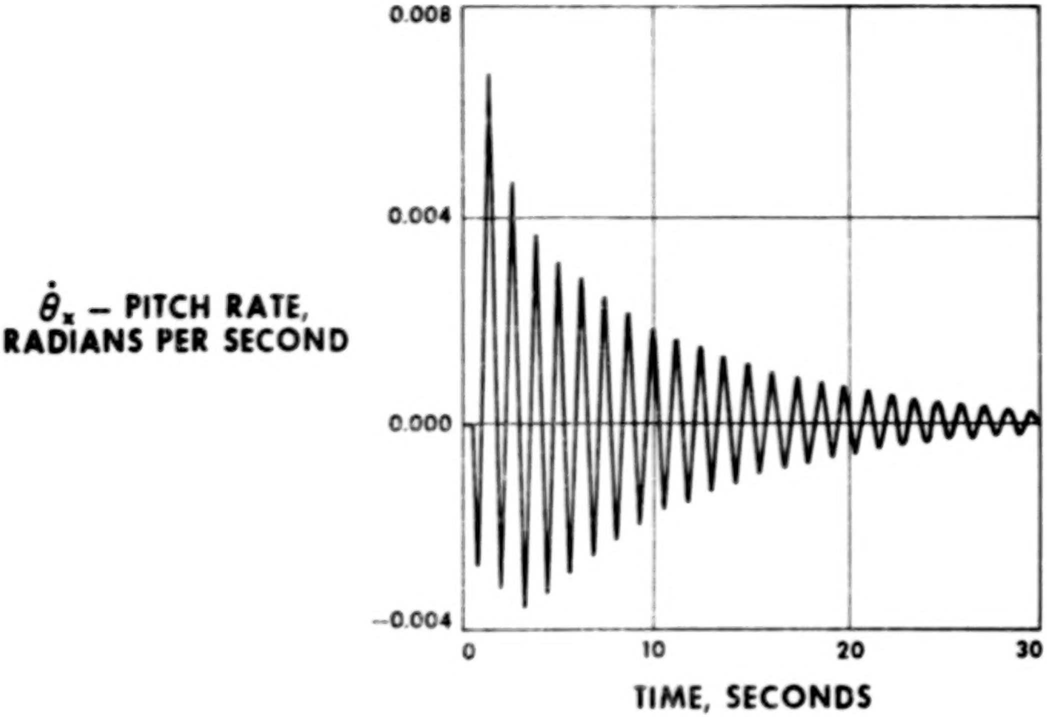


Figure 24.- Pitch rate vs time, flexible vehicle (gain = 0 dB)

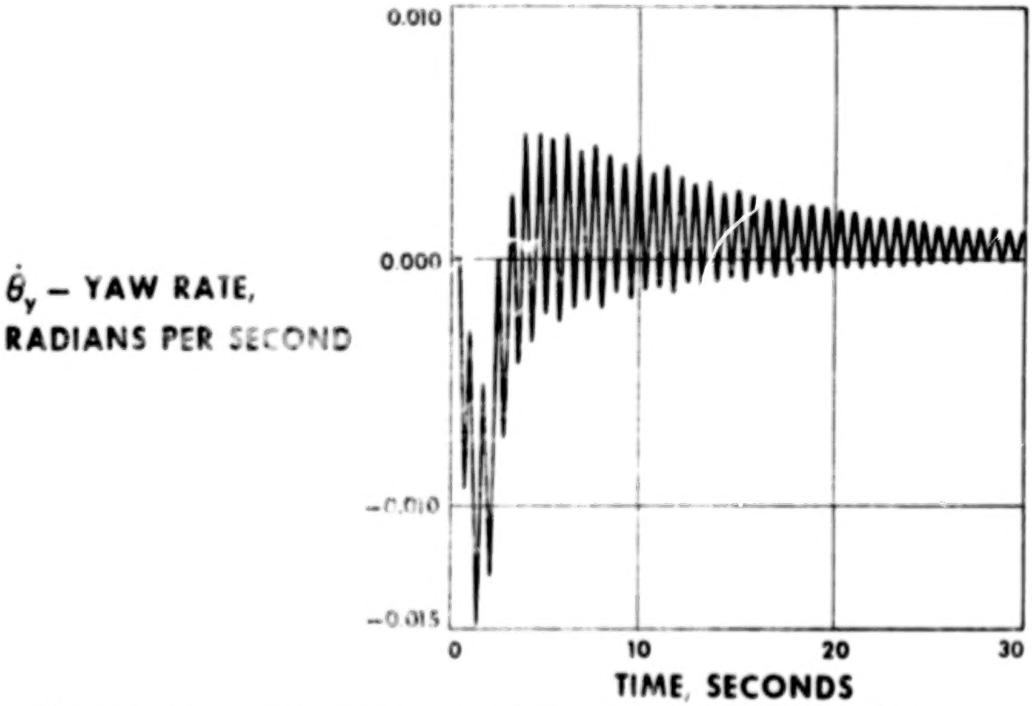


Figure 25.- Yaw rate vs time, flexible vehicle (gain = 0 dB)

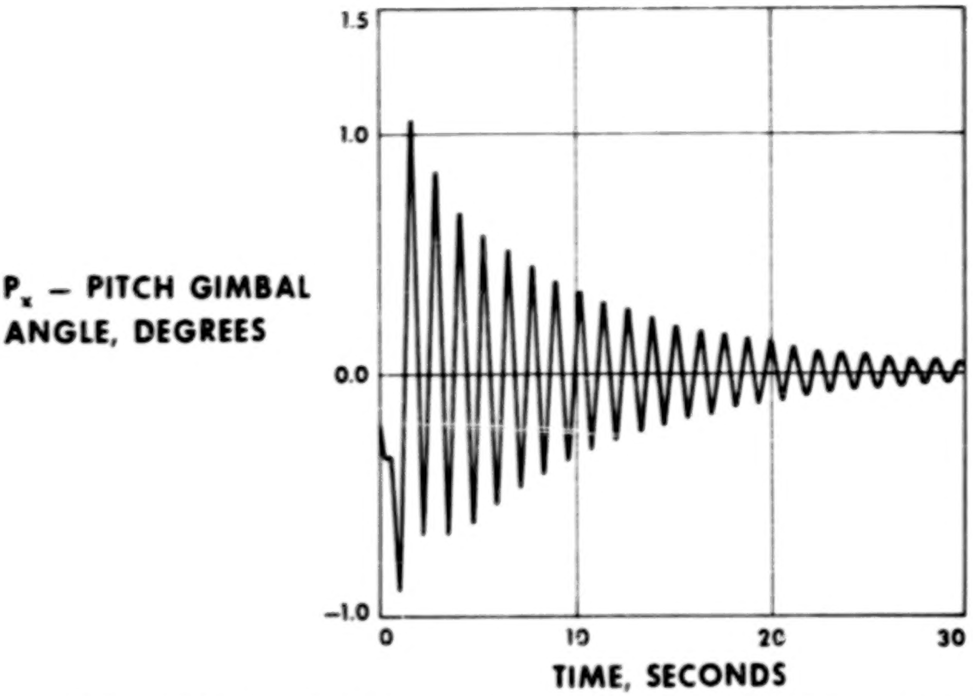


Figure 26.- Pitch gimbal angle vs time, flexible vehicle (gain = 0 dB)

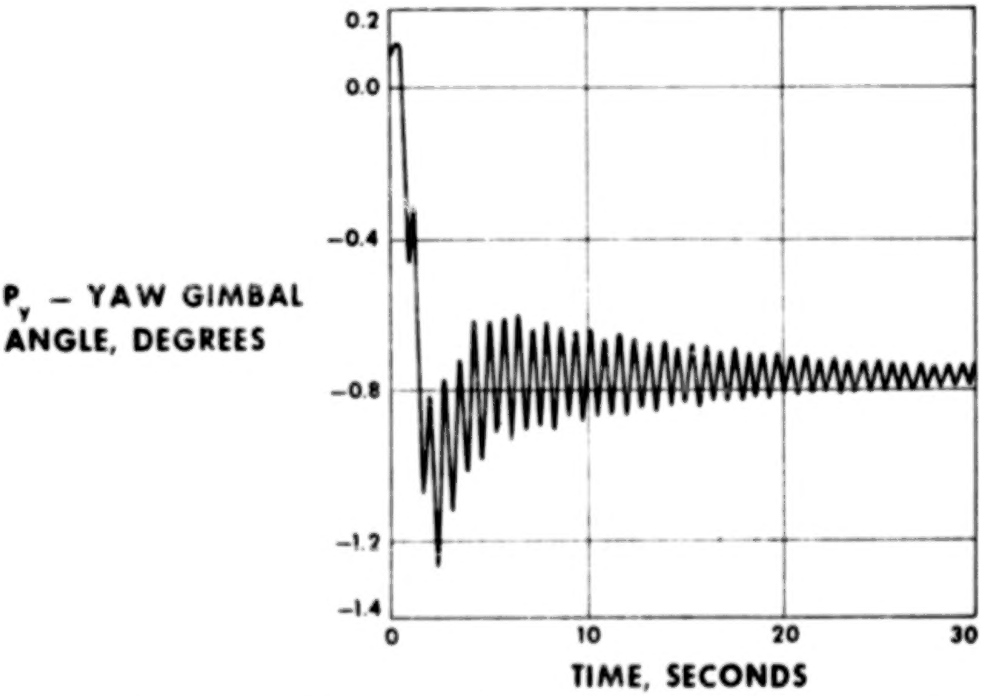


Figure 27.- Yaw gimbal angle vs time, flexible vehicle (gain = 0 dB)

attempts to reach some generalizations about the form of these loci that may be of help to the control designer. In addition, the interaxis coupling effect is being examined by an eigenvalue approach in the hope that this may also aid the designer and make unnecessary a heavy reliance on relatively expensive computer simulations.

REFERENCES

1. Kopf, E. H.: "A Mariner Orbiter Autopilot Design." Technical Report 32-1349, Jet Propulsion Laboratory, Pasadena, California, January 15, 1969.

2. Hooker, W. W., and Margulies, G.: "The Dynamical Attitude Equations for an n-Body Satellite." J. Astronaut. Sci., vol. 12, 1965, pp. 123-128.

3. Likins, P. W., and Wirsching, P. H.: "Use of Synthetic Modes in Hybrid Coordinate Dynamic Analysis." AIAA J., vol. 6, 1968, pp. 1867-1872.

4. Likins, P. W., and Gale, A. H.: "Analysis of Interactions Between Attitude Control Systems and Flexible Appendages." Paper presented at the 19th International Astronautical Congress, October 14-19, 1968.

5. Likins, P.W.: "Dynamics and Control of Flexible Space Vehicles." Technical Report 32-1329, Jet Propulsion Laboratory, Pasadena, California, February 15, 1969.

STEERING LAWS FOR DOUBLE-GIMBAL CONTROL-MOMENT GYROS

By Stephen W. Winder
Marshall Space Flight Center

SUMMARY

To generate accurate attitude and rate control in orbiting space vehicles, a continuous type of control is desired. One type of control actuator being considered to provide this continuous control is the control-moment gyro (CMG).

This paper is concerned with the derivation of CMG steering laws. The type of CMG considered is a double-gimbal one which has a constant magnitude momentum vector. The equations for the generation of torque of an inertially fixed CMG are presented. To realize a particular commanded torque, the gimbal rates of the CMG must be specified. These gimbal rates are shown to be functions of gimbal angle orientation and the commanded torque. The conventional cross-product steering law is derived by minimizing a function which is the square of the magnitude of the error between the commanded torque and the realizable torque of the gyro. This steering law is severely dependent upon the inner gimbal angle. When the inner gimbal reaches 90 degrees, the cross-product law requires infinite outer gimbal rates to realize finite torque values. By adding a gimbal rate penalty function to the original quantity to be minimized, a compromise is reached between control effort and torque-producing qualities. This compromise provides a control concept which removes the singularity of the cross-product law by requiring the outer gimbal rate to approach zero as the inner gimbal angle approaches 90 degrees.

When considering more than one gyro, a perfect matching of gyro torque to command torque is possible. However, no unique set of gimbal rates exist. In order to obtain a unique set of gimbal rates, and to ensure that these rates will not be too large, a gimbal rate constraint is introduced. This constraint has the same form as the rate penalty considered with one gyro. By use of the constraint relationship, steering laws are developed which yield unique solutions. The form of the cross-product law for more than one gyro is shown to be a special case of these steering laws.

By formulating the control concepts of the CMG's in a manner different than had been done previously, additional insight was gained into the mechanisms of CMG steering laws.

INTRODUCTION

The precision attitude control of orbiting spacecraft has produced a need for a continuous type of control rather than the usual reaction control jet system (ref. 1). The double-gimbal control-moment gyro (CMG) has been developed to fulfill this need. To use these CMG's for attitude control, a control law must be developed for the CMG's themselves. This control law, usually designated "steering law" to distinguish it from a general vehicle control law, is the subject of this paper. The original steering law for these gyros (ref. 2) was a result of physical reasoning from torque-producing qualities of the gyro. A more sophisticated law was developed called the "cross-product" law (ref. 3), which used a vector approach of torque generation to evolve a steering law. Recently, another formulation has been introduced (ref 4), which generates the steering laws by considering the torque equation of the gyro in a vector-matrix form. By the use of this formulation, insight may be gained into the development of steering laws, including the necessary gimbal rate constraints that must be added to ensure proper control.

LIST OF SYMBOLS

Symbol	Definition
ω_s	spin rate of CMG rotor
X,Y,Z	control coordinate system
X_m,Y_m,Z_m	mounting frame of the CMG
C	coordinate transformation (3x3) between the mounting frame and the control coordinate system
$\underline{\ }$	a solid underline indicates a vector quantity
δ_1,δ_3	inner and outer gimbal angles
$\underline{H}$	momentum vector of the CMG in mounting frame coordinates
H	magnitude of the momentum vector
$\underline{H}_c$	momentum of the CMG in the control frame coordinates
$\underline{H}_I$	momentum of the CMG expressed in inertial space coordinates
$\underline{T}_{out}$	output torque (3x1) vector of the CMG in inertial space with components $[T_{xout},T_{yout},T_{zout}]$

.	dot over a quantity indicates a time derivative
D	matrix (3x2) relating the gimbal rates to the output torques
$[\ ]^t$	superscript t indicates the transpose of a matrix
$[\ ]^{-1}$	inverse of a matrix
Q	positive definite matrix (2x2)
k	a scalar gain constant
$\left\{\begin{array}{l} H_i, C_i, D_i \\ \underline{\delta}(i), \underline{T}_{out}(i) \\ \underline{H}(i) \end{array}\right.$	those quantities previously defined but here applied to the ith CMG of an n gyro distribution
$\overset{*}{D}$	dynamic matrix (3x2n) for an n gyro distribution
$\underline{\overset{*}{\delta}}^t$	gimbal angle (1x2n) vector for an n gyro distribution
$\overset{*}{Q}$	positive definite matrix (2nx2n)
$\underline{\omega}_{CI}$	rotational velocity of the control frame with respect to an inertial frame
$\underline{T}_{com}$	the commanded torque in the control frame with components $[T_{xcom}, T_{ycom}, T_{zcom}]$
$\hat{\underline{T}}_{com}$	the commanded torque resolved in the CMG mounting frame with components $[\hat{T}_{xcom}, \hat{T}_{ycom}, \hat{T}_{zcom}]$

STEERING LAW EQUATIONS FOR ONE DOUBLE-GIMBAL CMG

A model of the double-gimbal CMG of the type currently proposed for the NASA-AAP program is shown in Figure 1. The spin rate ω_s of the rotor is held constant so that the momentum vector $\underline{H}$ of the gyro has a constant magnitude H. The various coordinate systems in the figure are associated with the gimbal rotations δ_1 and δ_3 and the relationship between the mounting frame for the gyro and the control frame of the body to be controlled by the gyro. The momentum vector of the gyro resolved in the control frame is:

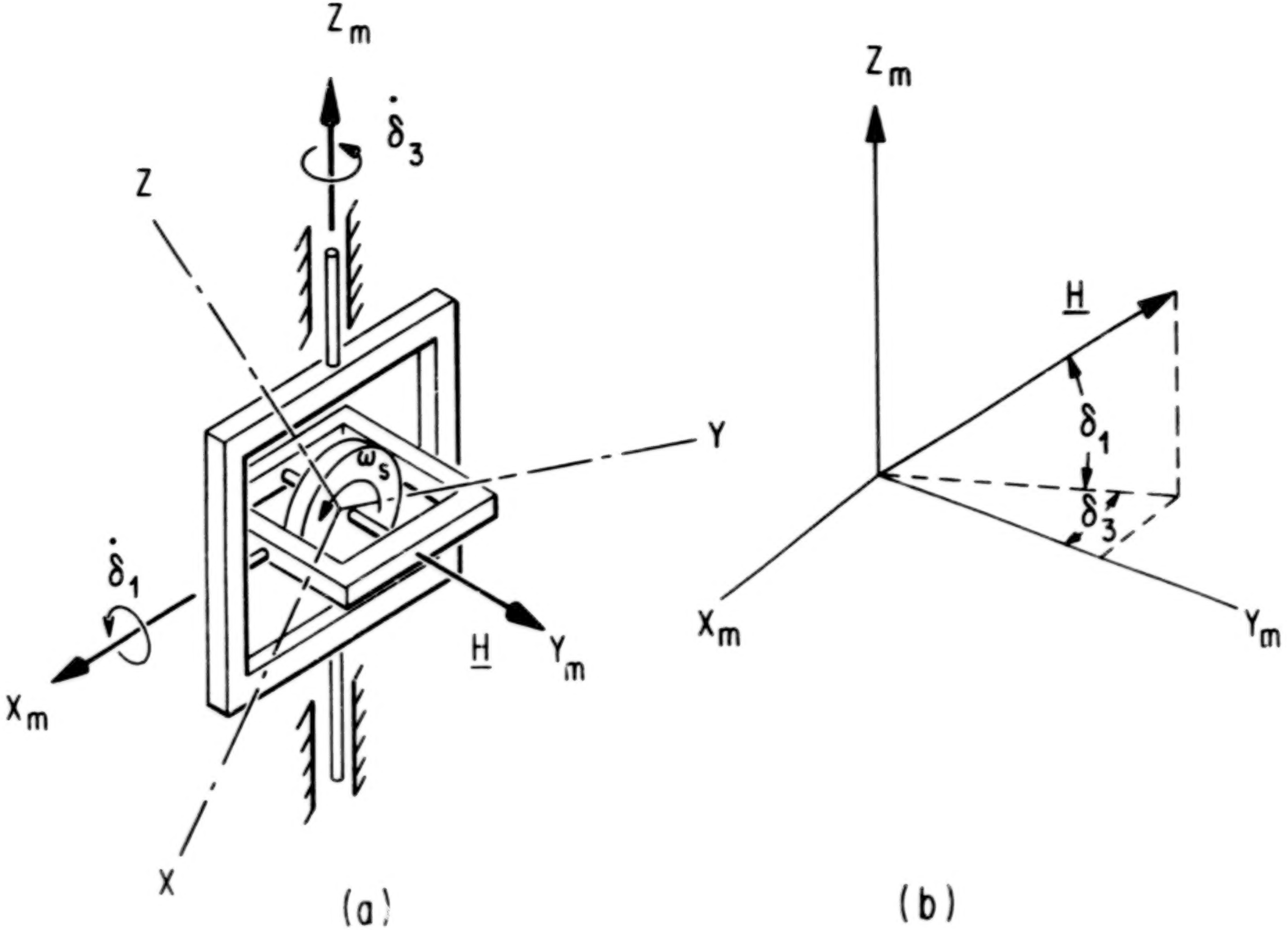

Z_m
$\dot{\delta}_3$
Z
Y
ω_s
$\dot{\delta}_1$
X_m
H
Y_m
X
(a)
Z_m
H
δ_1
δ_3
X_m
Y_m
(b)

Figure 1.- (a) Double gimbal CMG (X_m, Y_m, Z_m) mounting frame (X, Y, Z) control frame; (b) location of H vector for gimbal angle rotations wrt. to mounting frame X_m, Y_m, Z_m

$$\underline{H}_C = C\underline{H} = HC \begin{bmatrix} \cos\delta_1 \cos\delta_3 \\ -\cos\delta_1 \sin\delta_3 \\ -\sin\delta_1 \end{bmatrix} \qquad (1)$$

where C is a 3x3 constant matrix which relates the orientation of the gyro to the control frame.

The torque that may be generated from this gyro is found by expressing the time rate of change of the gyro momentum in inertial space (ref. 5):

$$\underline{T}_{out} = \dot{\underline{H}}_I = \dot{\underline{H}}_C + \underline{\omega}_{CI} \times \underline{H}_C \qquad (2)$$

In most applications where this type of gyro is used, the control frame is to be held in an inertial frame so that $\underline{\omega}_{CI}$ is very small. When the required differentiation is performed and the $\underline{\omega}_{CI} \times \underline{H}_C$ term neglected, the output torque of the gyro becomes:

$$\underline{T}_{out} \doteq \dot{\underline{H}}_C = HCD\dot{\underline{\delta}} \quad . \qquad (3)$$

The matrix D is a 3x2 matrix of the form

$$D = \begin{bmatrix} -\sin\delta_1 \cos\delta_3 & -\cos\delta_1 \sin\delta_3 \\ \sin\delta_1 \sin\delta_3 & -\cos\delta_1 \cos\delta_3 \\ -\cos\delta_1 & 0 \end{bmatrix}$$

and $\dot{\underline{\delta}}$ is a 2x1 vector of gimbal rates $\dot{\delta}_1$ and $\dot{\delta}_3$. Since the purpose of the CMG is to produce torques in the control frame, the control vector $\dot{\underline{\delta}}$ must be generated so that the gyro produces the commanded torque. Thus, we are led to the notion of finding a relationship between the torque command $\underline{T}_{com}$ and the gimbal rates $\dot{\underline{\delta}}$ that will produce the commanded torque from the gyro (Figure 2). An assumption implicit in this development is that any gimbal rate may be achieved instantaneously. While this is not strictly true, the motors that do rotate the gimbals are very fast acting.

Since the rank of the D matrix is, at most, two, the output torque of the CMG cannot completely span a three-dimensional torque space. A realization of an arbitrary three-dimensional torque is then impossible. The physical significance of this statement may be obtained from Figure 3. For an arbitrary orientation of the momentum vector, the output torque will lie in a plane normal to the momentum vector. This plane is, of course, the two-dimensional space spanned by the matrix D. Unless the commanded torque lies in this plane, a perfect realization will not be possible. For a general torque $\underline{T}_{com}$, the best that may be hoped for is to produce an output torque which is the projection of the commanded torque on the plane. A precise way to generate this torque is to choose $\dot{\underline{\delta}}$ so as to minimize the square of the magnitude of the error between the command torque and the output torque of the gyro. This minimization amounts to an instantaneous optimization which is dependent on the gimbal angles as well as the gimbal rates:

$$(\underline{T}_{com}-\underline{T}_{out})^t\,(\underline{T}_{com}-\underline{T}_{out}) = \underline{T}^t_{com}\,\underline{T}_{com} - 2H\underline{T}^t_{com}\,CD\dot{\underline{\delta}} + H^2\dot{\underline{\delta}}^t\,D^t\,D\dot{\underline{\delta}} \tag{4}$$

Minimizing Eq. (4) with respect to $\dot{\underline{\delta}}$ yields:

$$\dot{\underline{\delta}} = \frac{1}{H}\,(D^tD)^{-1}\,D^t\,C\underline{T}_{com} \tag{5a}$$

For $\hat{\underline{T}}_{com} = C\underline{T}_{com}$, Eq. (5a) may be written as

$$\begin{bmatrix} \dot{\delta}_1 \\ \\ \dot{\delta}_3 \end{bmatrix} = \frac{1}{H}\begin{bmatrix} -\sin\delta_1\cos\delta_3 & \sin\delta_1\sin\delta_3 & -\cos\delta_1 \\ \\ -\sec\delta_1\sin\delta_3 & -\sec\delta_1\cos\delta_3 & 0 \end{bmatrix}\begin{bmatrix} \hat{T}_{xcom} \\ \hat{T}_{ycom} \\ \hat{T}_{zcom} \end{bmatrix} \tag{5b}$$

This law is called the "cross-product" steering law (ref. 3). Originally it was derived by defining a velocity vector by taking the cross-product of the momentum vector and the command torque. The components of this velocity vector were then equated to the gimbal rates resolved in the same coordinate frame.

The 2x2 matrix:

$$D^tD = \begin{bmatrix} 1 & 0 \\ \\ 0 & \cos^2\delta_1 \end{bmatrix}$$

is seen to possess a singularity when the inner gimbal angle reaches 90 degrees. This results in an extremely high gimbal rate for the outer gimbal to realize small finite torques. A physical interpretation of the phenomenon is shown in Figure 4 where the view is down the inner gimbal axis. The outer gimbal rate δ_3 is capable of producing torques perpendicular to the paper. The magnitude of the torque is proportional to the cosine of the inner gimbal angle. When δ_1 is very near 90 degrees, a very large gimbal rate is required to produce a finite torque. This statement is just a way of saying that a large amount of control effort is expended for very little benefit. This condition is commonly called "gimbal lock." Realizing that the secant effect is detrimental to the operation of the CMG, we approximated the secant in Eq. (5b) by a constant, nominally 1, to yield a modified cross-product steering law (ref. 3):

$$\begin{bmatrix} \dot{\delta}_1 \\ \\ \dot{\delta}_3 \end{bmatrix} = \frac{1}{H} \begin{bmatrix} -\sin \delta_1 \cos \delta_3 & \sin \delta_1 \sin \delta_3 & -\cos \delta_1 \\ \\ -\sin \delta_3 & -\cos \delta_3 & 0 \end{bmatrix} \begin{bmatrix} \hat{T}_{xcom} \\ \hat{T}_{ycom} \\ \hat{T}_{zcom} \end{bmatrix} \tag{6}$$

This replacement of the secant by a constant does prevent the generation of excessive gimbal rates for large inner gimbal angles, but the outer gimbal axis is allowed to rotate at a finite rate, even though the torque output is ineffectual. In an effort to reduce the problem of excessive gimbal rates, the torque-fitting criterion [Eq. (4)] is altered to include a penalty function on the gimbal rates. This penalty function has the form $\underline{\dot{\delta}}^t Q \underline{\dot{\delta}}$ where Q is a positive definite matrix. The addition of this element reflects a balancing between the attempted torque fit and the amount of control effort required to accomplish the task. Minimizing this quantity with respect to the gimbal rates gives:

$$\underline{\dot{\delta}} = \frac{1}{H} (HQ + D^t D)^{-1} D^t \underline{\hat{T}}_{com} \tag{7}$$

The inverse of the matrix will always exist because it is the sum of a positive definite matrix and a positive semidefinite matrix. Since $(HQ + D^t D)^{-1}$ always exists, the main distributional effect is shifted to the matrix D^t. For this reason, it is called the "transpose" steering law:

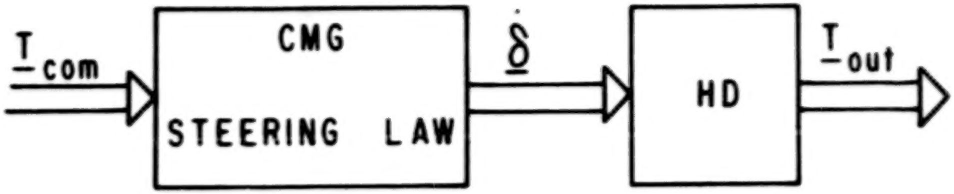

$\underline{T}_{com}$
CMG
STEERING LAW
$\dot{\underline{\delta}}$
HD
$\underline{T}_{out}$

Figure 2.- CMG steering law block diagram representation

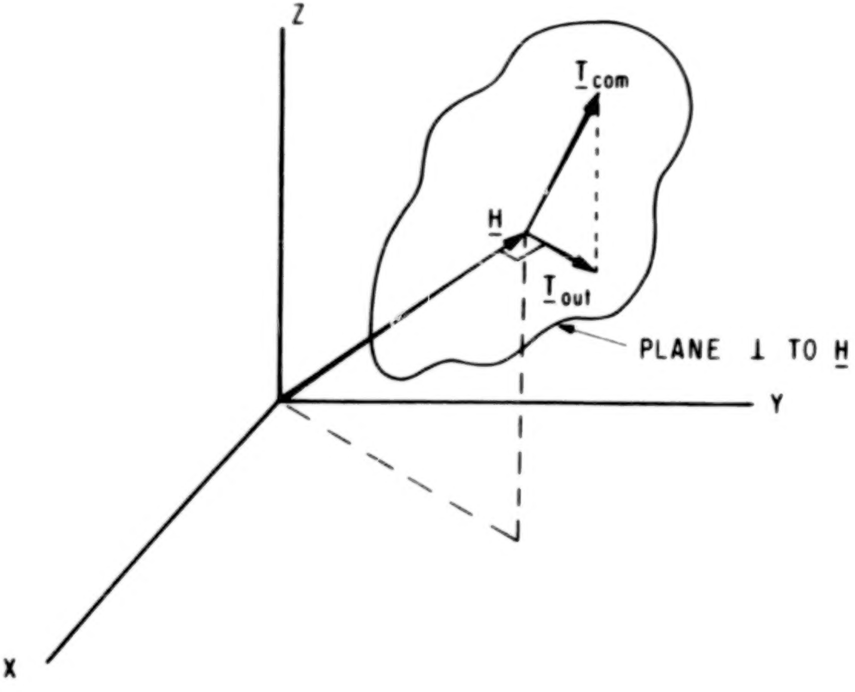

Z
$\underline{T}_{com}$
$\underline{H}$
$\underline{T}_{out}$
PLANE ⊥ TO $\underline{H}$
Y
X

Figure 3.- Relationship between commanded torque and output torque of a double gimbal gyro

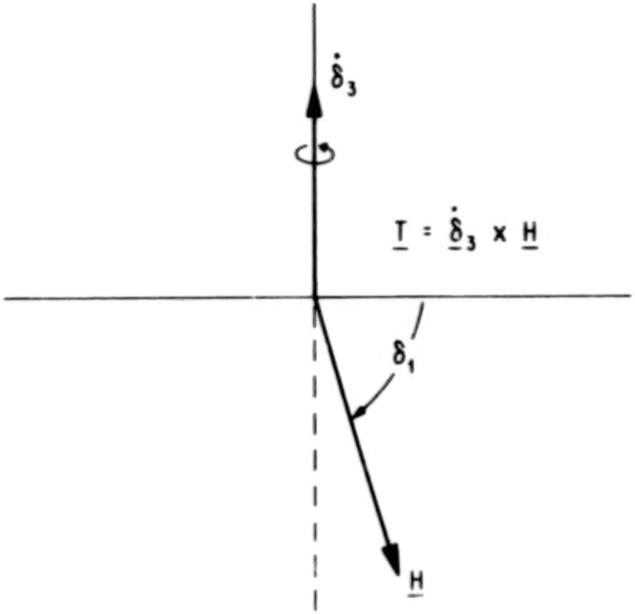

$\dot{\delta}_3$
$\underline{T} = \dot{\underline{\delta}}_3 \times \underline{H}$
δ_1
$\underline{H}$

Figure 4.- Relationship between outer gimbal rate and the momentum vector

$$\begin{bmatrix} \dot{\delta}_1 \\ \\ \dot{\delta}_3 \end{bmatrix} = \frac{1}{H}\,(HQ + D^tD)^{-1}$$

$$\begin{bmatrix} -\sin\delta_1\cos\delta_3 & \sin\delta_1\sin\delta_3 & -\cos\delta_1 \\ \\ -\cos\delta_1\sin\delta_3 & -\cos\delta_1\cos\delta_3 & 0 \end{bmatrix} \begin{bmatrix} \hat{T}_{xcom} \\ \hat{T}_{ycom} \\ \hat{T}_{zcom} \end{bmatrix} \qquad (8)$$

the last row of the matrix now includes the cosine of the inner gimbal angle rather than the secant or a constant. The outer gimbal rate approaches zero rather than infinity or a finite value when the inner gimbal angle approaches 90 degrees. A simple example will illustrate the differences between the cross-product law, the modified cross-product law, and the transpose law.

Example

Initially the inner and outer gimbal angles are set to zero, and the matrix C is set equal to the identity matrix. The momentum vector, which is normalized to one, is pointing down the x axis. A torque is desired along the y axis. The command torque is:

$$\underline{T}_{com} = \begin{bmatrix} 0 \\ \underline{T}_{ycom} \\ 0 \end{bmatrix}$$

The inner gimbal rates for each law are a function of the same row vector, and for this example are zero for any position of the inner gimbal angle provided the outer gimbal angle remains at zero. The inner gimbal will be varied from zero to 90 degrees, while the outer gimbal angle remains at zero. The gimbal rates and torques for each law are:

Regular Cross-Product

$\dot{\delta}_1 = 0$ $\qquad T_{xout} = 0$ $\qquad T_{zout} = 0$

$\dot{\delta}_3 = \sec\delta_1\, T_{ycom}$ $\qquad T_{yout} = T_{ycom}$

Modified Cross-Product

$$\dot{\delta}_1 = 0 \qquad T_{xout} = 0 \qquad T_{zout} = 0$$

$$\dot{\delta}_3 = T_{ycom} \qquad T_{yout} = \cos\delta_1 \, T_{ycom}$$

Transpose Law with $Q = kI$

$$\dot{\delta}_1 = 0 \qquad T_{xcom} = 0 \qquad T_{zout} = 0$$

$$\dot{\delta}_3 = \frac{\cos\delta_1}{(k+\cos^2\delta_1)} T_{ycom} \qquad T_{yout} = (k+\cos^2\delta_1)^{-1}\cos^2\delta_1 \, T_{ycom}$$

These rates and torques are plotted in Figures 5 and 6. The choice of Q equal to kI was made because each gimbal rate is controlled by its individual torque motor. The value of k was introduced to investigate how the penalty affects the output torques. The graphs demonstrate the trade-offs between the various steering laws which are summarized as follows:

1. The torque gain of the regular cross-product law is unity, but its outer gimbal rate approaches infinity as the inner gimbal angle approaches 90 degrees.
2. The torque gain of the modified cross-product law goes to zero as the inner gimbal approaches 90 degrees, but the gimbal rate remains constant even in the gimbal lock position.
3. For small values of k, the rate and torque profiles are similar to the cross-product law rates and torques except in the region near gimbal lock. In this position, both the rate and torque go to zero. As the value of k is increased, the rate is reduced, but so is the torque output. The value of k is an adjustable parameter that may be varied according to the constraints of the problem at hand. From the graphs, it is evident that the value of k may be used to obtain a good working compromise of each law.

Another interesting feature of the rate penalty formulation is the implementation of the law. As opposed to the other two laws, the transpose law may be realized in a feedback fashion (Figure 7). The concept of torque feedback enters naturally as a result of rate penalization.

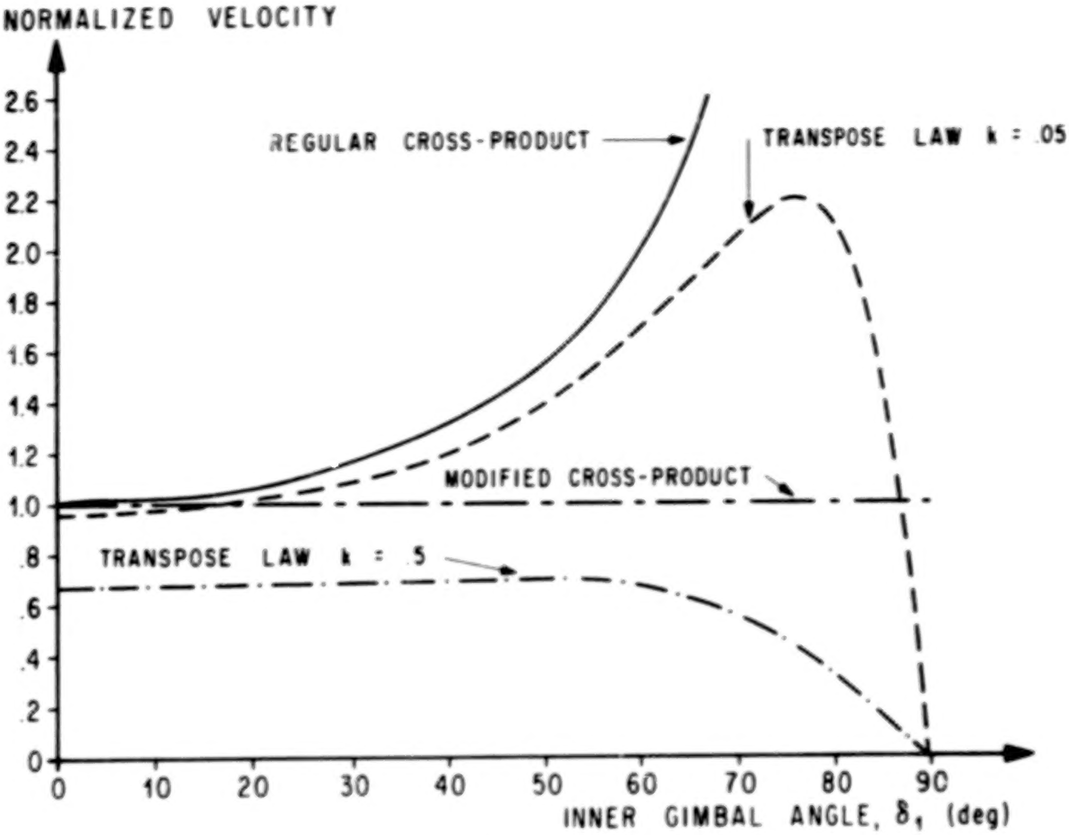


Figure 5

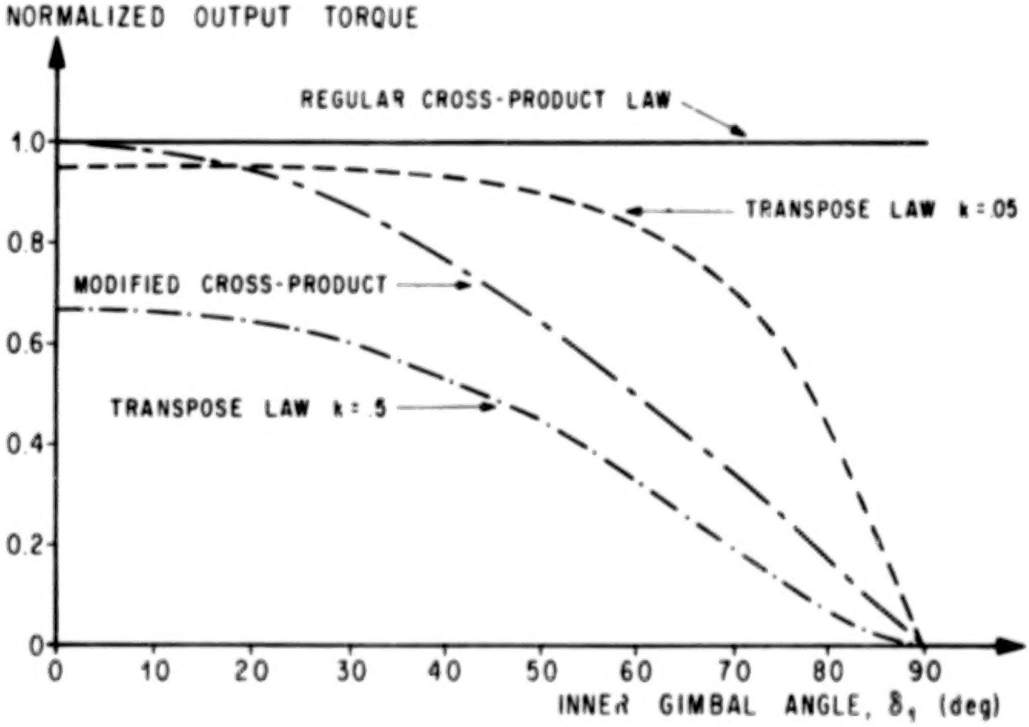


Figure 6

Steering Laws for n Double Gimbal CMG's

In this section, the concepts treated previously will be used for a distribution of n CMG's. The output torque produced by the i-th gyro will be written as

$$\underline{T}_{out}(i) = H_i C_i D_i \underline{\delta}(i) \tag{9}$$

where all quantities have been defined previously. The total torque of the gyro distribution is the vector sum of all the gyro torques:

$$\underline{T}_{out} = \sum_{i=1}^{n} H_i C_i D_i \dot{\underline{\delta}}(i) \tag{10}$$

Equation (10) may be rewritten in matrix form as

$$\underline{T}_{out} = \overset{\star}{D} \dot{\overset{\star}{\underline{\delta}}}(i) \tag{11}$$

where $\overset{\star}{D}$ is a 3x2n matrix = $(H_1 C_1 D_1 \mid H_2 C_2 D_2 \mid \ldots \mid H_n C_n D_n)$

and $\overset{\star}{\underline{\delta}}^t$ is a 1x2n vector $(\underline{\delta}^t(1) \mid \underline{\delta}^t(2) \mid \ldots \mid \underline{\delta}^t(n))$

When the value of n is 2 or greater, a perfect realization of the command is possible if the matrix $\overset{\star}{D}$ is at its maximum rank of 3. However, because there are more degrees of freedom than constraints, there is an infinite number of gimbal rates that will satisfy Eq. (11) such that $\underline{T}_{out} = \underline{T}_{com}$. No guarantee is available to ensure that the gimbal rates will always stay within appropriate bounds. One approach to ensure reasonable and unique gimbal rates is to minimize

$$\dot{\overset{\star}{\underline{\delta}}}^t \, \overset{\star}{Q} \, \dot{\overset{\star}{\underline{\delta}}} \tag{12a}$$

subject to

$$\overset{\star}{D} \dot{\overset{\star}{\underline{\delta}}} - \underline{T}_{com} = 0 \tag{12b}$$

where $\overset{*}{Q}$ is a 2nx2n positive definite matrix. A two-dimensional analog to this formulation is described in Figure 8 where $\overset{*}{D}$ is a 1x2 vector (d_1,d_2), $\underline{T}_{com}$ is a scalar, and $\dot{\delta}$ is a 2x1 vector. The straight line is the constraint relationship, and the circle or ellipse the quantity to be minimized. Since any point on the line satisfies the constraint, one may visualize how high gimbal rates could occur. The desired solution is the point of tangency between the two figures. Solving the indicated minimization yields the gimbal rates:

$$\underline{\dot{\overset{*}{\delta}}} = \overset{*}{Q}^{-1} \overset{*}{D}^{t} (\overset{*}{D}\overset{*}{Q}^{-1} \overset{*}{D}^{t})^{-1} \underline{T}_{com} \tag{13}$$

If $\overset{*}{Q} = kI$, the gimbal rate equation reduces to

$$\underline{\dot{\overset{*}{\delta}}} = \overset{*}{D}^{t} (\overset{*}{D}\overset{*}{D}^{t})^{-1} \underline{T}_{com} \tag{14}$$

which is independent of k.

The question of the rank of $\overset{*}{D}$ is still under investigation. One case when $\overset{*}{D}$ has rank 2 is easily recognizable. This occurs when the momentum vectors $\underline{H}(i)$ are all collinear. To use this law, the collinearity of the vectors must be avoided, or an alternate law must be used when these situations are encountered. A second approach is to use the quadratic criterion:

$$(\underline{T}_{com} - \underline{T}_{out})^{t} (\underline{T}_{com} - \underline{T}_{out}) + \dot{\overset{*}{\delta}}^{t} \overset{*}{Q} \underline{\dot{\overset{*}{\delta}}} \tag{15}$$

Minimizing with respect to $\underline{\dot{\overset{*}{\delta}}}$ yields:

$$\underline{\dot{\overset{*}{\delta}}} = (\overset{*}{Q} + \overset{*}{D}^{t}\overset{*}{D})^{-1} \overset{*}{D}^{t} \underline{T}_{com} \tag{16}$$

The matrix $(\overset{*}{Q} + \overset{*}{D}^{t}\overset{*}{D})$ is positive definite, and the solution to Eq. (16) will always exist and be unique. If $\overset{*}{Q} = \frac{1}{k} I$, the solution to Eq. (16) may be realized by the torque feedback shown in Figure 9. The value of k may be adjusted in much the same way as was done when one gyro was considered. Probably the strongest feature of this formulation is the simplicity of implementation where a feedback path replaces a continuous matrix inversion [Eq. (13)]. Notice also that the transpose concept is preserved for each individual gyro. If the $\overset{*}{Q}$ matrix is of the form

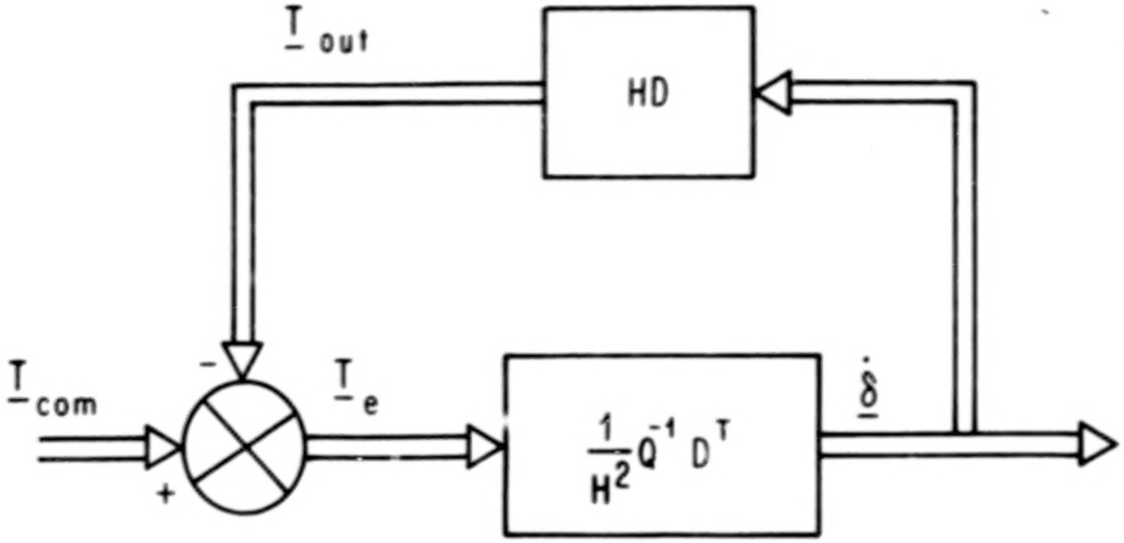

$\underline{T}_{out}$
HD
$\underline{T}_{com}$
$-$
$+$
$\underline{T}_e$
$\frac{1}{H^2} Q^{-1} D^T$
$\dot{\underline{\delta}}$

Figure 7.- Torque feedback realization

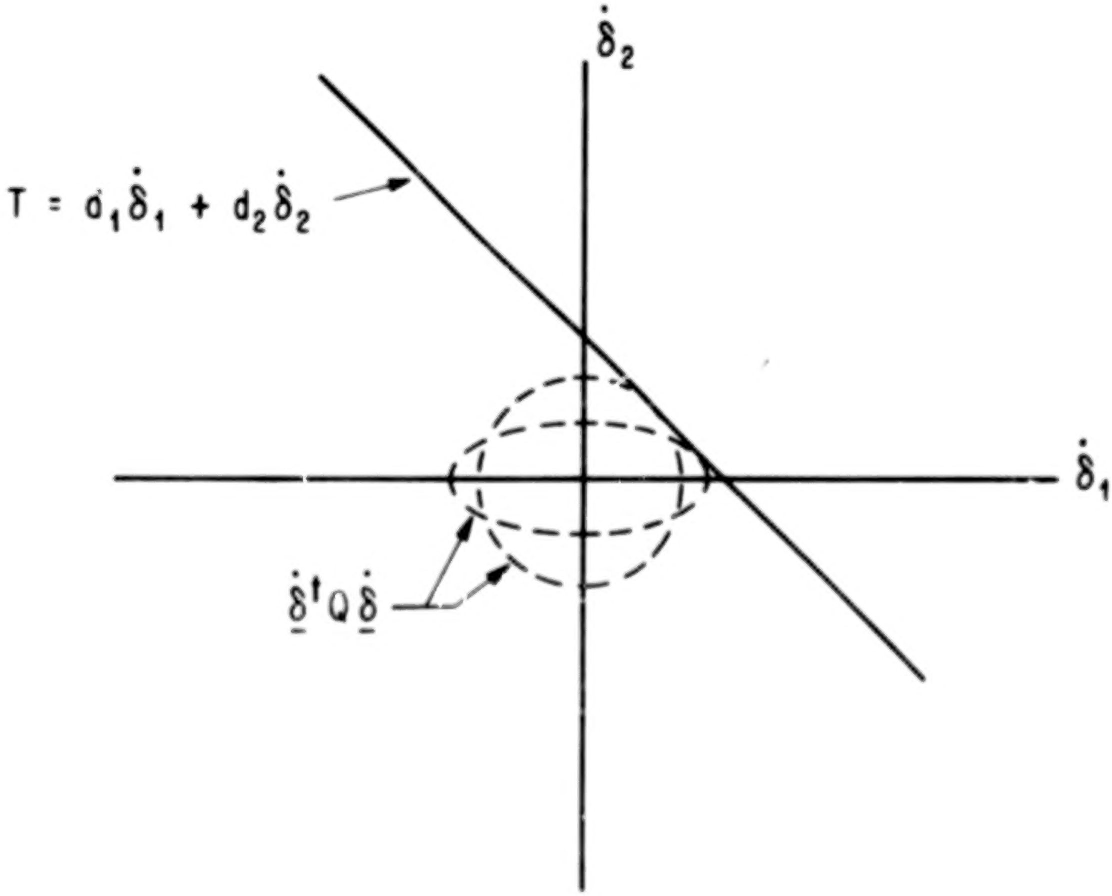

$\dot{\delta}_2$
$T = d_1\dot{\delta}_1 + d_2\dot{\delta}_2$
$\dot{\delta}_1$
$\dot{\underline{\delta}}^t Q \dot{\underline{\delta}}$

Figure 8.- Two-dimensional representation of Eqs. (12a) and (12b)

TABLE OF CONTENTS

1. SPACE VEHICLE CONTROL APPLICATIONS

Design Development of the Apollo Lunar Module
Kenneth J. Cox
Manned Spacecraft Center 1 1/A10

Design Development of the Apollo Command and Service Module Thrust Vector Attitude Control System
William H. Peters
Manned Spacecraft Center 63 1/E3

Orbiting Astronomical Observatory (OAO-A2) Stabilization and Control Subsystem
Thomas E. Huber
Goddard Space Flight Center 83 1/G10

Applications of the Hybrid Coordinate Method to the TOPS Autopilot
G. E. Fleischer
Jet Propulsion Laboratory 127 2/D7

Steering Laws for Double-Gimbal Control-Moment Gyros
Stephen W. Winder
Marshall Space Flight Center 157 2/F13

2. AIRCRAFT GUIDANCE, CONTROL AND HANDLING QUALITIES

Optimum Horizon Guidance Techniques for Aircraft
Heinz Erzberger and Homer Q. Lee
Ames Research Center 175 3/A8

Inertially-Aided, All-Weather Landing Systems for C/TOL.
Robert J. Pawlak
Electronics Research Center 185 3/B4

An Approach to the Determination of Aircraft Handling Qualities Using Pilot Transfer Functions
James J. Adams and Howard G. Hatch, Jr.
Langley Research Center 199 3/C4

A Frequency Domain Approach to Handling Qualities Design
William A. Wolovich
Electronics Research Center 217 3/D8

3. SYSTEM SIMULATION AND IDENTIFICATION

Optimal Regulator or Conventional? Setup Techniques for a Model Following Simulator Control System
Dwain A. Deets
Flight Research Center 237 3/F1

TABLE OF CONTENTS (CONT'D)

3. SYSTEM SIMULATION AND IDENTIFICATION (CONT'D)

Quasi-Optimal Control of a Moving-Base Simulator
Elwood C. Stewart
Ames Research Center 253 3/G3

Some Modern Control Techniques for Human Operator Modeling and Identification
Richard S. Shirley
Electronics Research Center 261 3/G11

4. ENGINE CONTROL AND ADVANCED PROPULSION TECHNIQUES

Normal Shock and Restart Controls for a Supersonic Airbreathing Propulsion System
George H. Neiner, Gary L. Cole, and
Francis J. Paulovich
Lewis Research Center 299 4/D2

Optimal Control of a Supersonic Inlet to Minimize Frequency of Inlet Upstart
Bruce Lehtinen, John R. Zeller, and
Lucille C. Geyser
Lewis Research Center 323 4/F1

Control System Studies for Thermionic Reactors
Richard J. Hermsen,
California State Polytechnic College
Henrik G. Gronroos
JPL, California Institute of Technology 337 4/G1

Application of Fluidics to New Control Components
Miles O. Dustin, Vernon D. Gebben, and
Robert E. Wallhagen
Lewis Research Center 365 5/B7

Minimum Energy Control of a Class of Electrically Driven Vehicles
Yilmaz E. Sahinkaya
JPL, California Institute of Technology 387 5/D6

5. ADVANCED CONTROL TECHNIQUES

A New Formulation for the Epsilon Method Applied to the Minimum Time-to-Climb Problem
Lawrence W. Taylor, Jr., and Harriet J. Smith
Flight Research Center 423 5/G1

TABLE OF CONTENTS (CONT'D)

5. ADVANCED CONTROL TECHNIQUES (CONT'D)

Determination of Optimal Control Using Imbedding of the Terminal Conditions
Raymond C. Montgomery
Langley Research Center **435** 5/G13

Stability of Constant Gain Systems with Vector Feedback
George L. von Pragenau
Marshall Space Flight Center **443** 6/A10

Synthesis of Finite Settling Time Discrete Systems
James A. Gatlin
Goddard Space Flight Center **473** 6/C12

Design and Performance of Heart Assist or Artificial Heart Control Systems
John A. Webb, Jr., and Vernon D. Gebben
Lewis Research Center **495** 6/E1

$$\overset{*}{Q} = \frac{1}{k}\left|\begin{array}{cccc} D_1^t\ D_1 & 0 & \cdots & 0 \\ \hdashline 0 & D_2^t\ D_2 & \cdots & 0 \\ \hdashline \vdots & \vdots & \vdots & \vdots \\ \hdashline 0 & 0 & 0 & D_n^t\ D_n \end{array}\right|$$

$\overset{*}{Q}$ will remain a positive definite matrix as long as all the inner gimbal angles remain away from the gimbal lock position. A block diagram for the steering law for this value of $\overset{*}{Q}$ is shown in Figure 10. This form is the cross-product law for a distribution of n CMG's (ref. 3). In the original formulation, the feedback path was added to reduce the torque cross-coupling that resulted when each gyro was treated independently with its individual cross-product law. For a true rate penalty, this $\overset{*}{Q}$ is not a good choice because the penalty on the outer gimbal rate goes to zero if the inner gimbal of that particular gyro approaches the gimbal lock position. The derivation of the conventional cross-product law is referenced here to show how alternate formulations may lead to the same or comparable results.

CONCLUSION

A solution to the problem of generating CMG steering laws by a mathematical analysis of physical concepts has been attempted. The main objective has been to generate laws using a different approach from previous ones. When studying the torque-producing qualities of one gyro, a need for a rate penalty function was apparent. The steering law that resulted is physically plausible and mathematically interesting since it involves the notion of torque feedback rather than the open-loop controllers that result from a conventional formulation. The example demonstrated how the rate penalty may be weighted against the output torque to achieve a meaningful compromise between existing laws.

When more than one gyro is considered, a perfect fit of output torque to command torque is possible. However, a nonuniqueness of gimbal rates is present. To generate unique solutions and, at the same time, provide a meaningful constraint, a rate penalty was necessary. The steering laws that were evolved yielded unique gimbal rates, and in the case of the compromise between the minimization of the torque error squared and rate penalty, a feedback solution evolved. Also, by assuming a specific form of $\overset{*}{Q}$, the standard cross-product law was derived.

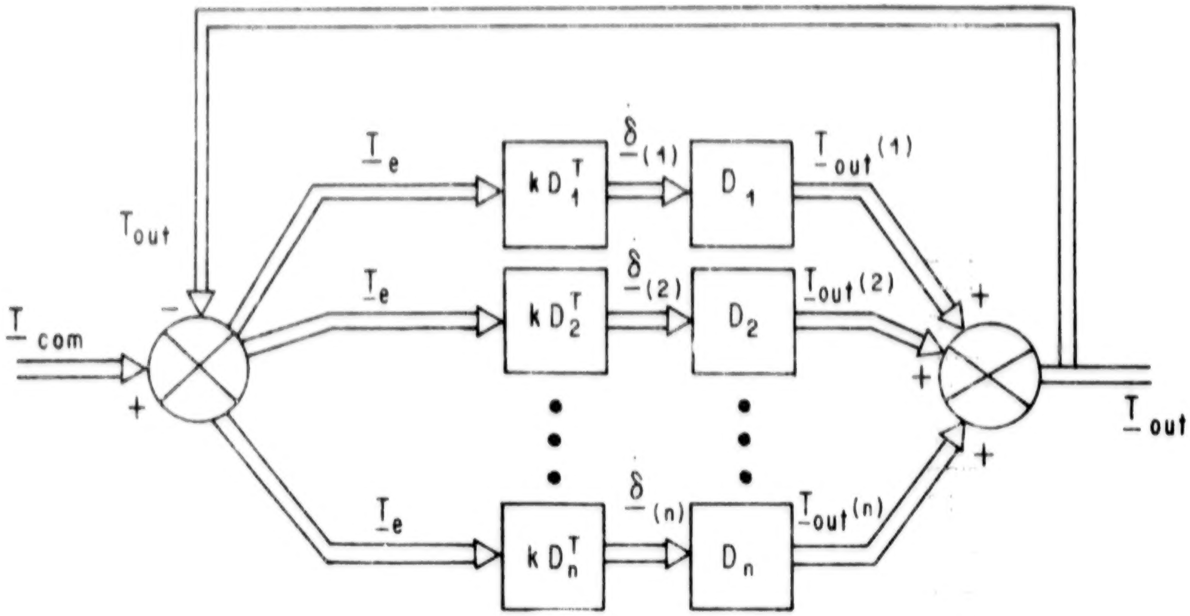

$\underline{T}_e$
$k D_1^T$
$\dot{\underline{\delta}}_{(1)}$
D_1
$\underline{T}_{out}^{(1)}$
T_{out}
$\underline{T}_e$
$k D_2^T$
$\dot{\underline{\delta}}_{(2)}$
D_2
$\underline{T}_{out}^{(2)}$
$\underline{T}_{com}$
$-$
$+$
$+$
$+$
$+$
$\underline{T}_{out}$
$\underline{T}_e$
$k D_n^T$
$\dot{\underline{\delta}}_{(n)}$
D_n
$\underline{T}_{out}^{(n)}$

Figure 9.- Feedback realization for n gyros

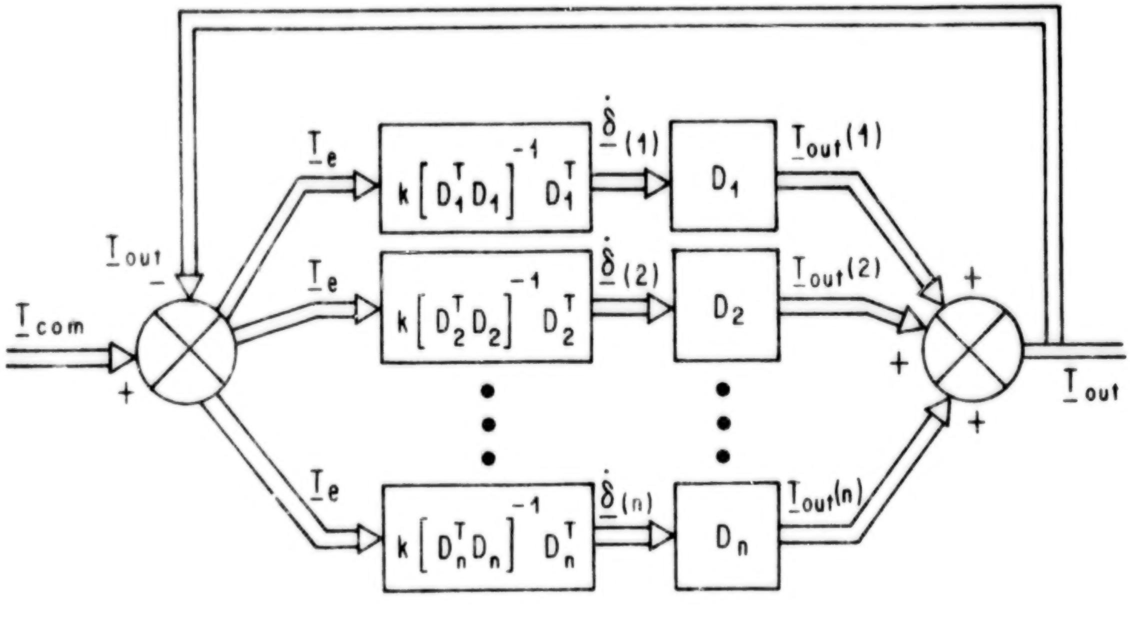

$\underline{T}_e$
$k\left[D_1^T D_1\right]^{-1} D_1^T$
$\dot{\underline{\delta}}_{(1)}$
D_1
$\underline{T}_{out}^{(1)}$
$\underline{T}_{out}$
$\underline{T}_e$
$k\left[D_2^T D_2\right]^{-1} D_2^T$
$\dot{\underline{\delta}}_{(2)}$
D_2
$\underline{T}_{out}^{(2)}$
$\underline{T}_{com}$
$-$
$+$
$+$
$+$
$+$
$\underline{T}_{out}$
$\underline{T}_e$
$k\left[D_n^T D_n\right]^{-1} D_n^T$
$\dot{\underline{\delta}}_{(n)}$
D_n
$\underline{T}_{out}^{(n)}$

Figure 10.- Conventional cross-product law

The advantage of the alternate formulation is evident in the additional insight that it provides about the torque-generating capabilities of the gyro or groups of gyros.

REFERENCES

1. Chubb, W. B.; and Epstein, M.: "Application of Control Moment Gyros in the Attitude Control of the Apollo Telescope Mount." AIAA Guidance, Control, and Flight Dynamics Conference, AIAA paper No. 68-866, Pasadena, California, August 12-14, 1968.
2. Kurzhals, Peter R.; and Grantham: "A System for Inertial Experiment Pointing and Attitude Control." NASA TR R-247, August 1966.
3. Morine, L. A.; and O'Connor, B. J.: "A Description of the CMG and Its Application to Space Vehicle Control." AIAA Guidance, Control, and Flight Dynamics Conference, AIAA paper No. 67-589, August 1967.
4. Ross, C. H.: "Antisingular Optimal Control of Control Moment Gyros." Northrop Corporation Report TR-795-9-548, May 1969.
5. Goldstein, H.: Classical Mechanics. Addison-Wesley Publishing Co., Inc., June 1959, p. 158.

Section 2

AIRCRAFT GUIDANCE CONTROL AND HANDLING QUALITIES

OPTIMUM HORIZONTAL GUIDANCE TECHNIQUES FOR AIRCRAFT

By Heinz Erzberger and Homer Q. Lee
Ames Research Center

SUMMARY

In the design of advanced flight control systems, consideration is currently being given to the problem of performing various terminal area maneuvers automatically. This paper discusses some problems of automatic guidance of an aircraft in the horizontal plane. The horizontal guidance tasks, which such a flight control system should be capable of performing, have been identified as being of three types: guiding the aircraft from any initial location and initial heading to (a) any final location and heading; (b) intercept and fly along a line of specified direction; (c) a final location with arbitrary final heading. Guidance problems such as capturing an ILS beam at a specified point on the beam, intercepting a VOR radial, and point-to-point navigation can be formulated in terms of these problems. The guidance laws given in this paper minimize the arc distance to fly between initial and final conditions subject to constraints on the minimum turning radius. Application of the Maximum Principle to these problems has shown that the optimum trajectories in problem (a) contain no more than four minimum radius turns or no more than two turns and one straight line segment. Trajectories in problem (b) contain no more than two turns and one straight line segment, and in problem (c) no more than one turn and a straight line segment. A technique for automatic terminal area guidance is briefly described.

INTRODUCTION

Ames Research Center is conducting a study of design concepts for advanced flight control systems. These systems have been referred to at Ames and elsewhere as pilot managed flight control systems to indicate that the pilot's role in them is seen chiefly as that of a manager or supervisor.

The main incentive for studying new concepts in flight control systems is to improve the flight safety and operational efficiency of advanced aircraft which must operate in a high density air traffic environment. The avenues through which these improvements can be reached are believed to be extended all weather landing and takeoff capability, a reduction of pilot workload, particularly in the critical phases of flight in and near the terminal area, and a reduction of air traffic controller workload. The implication of the last avenue mentioned is that advanced flight control systems may have to be designed to serve not only the flight crew, as is traditional, but also the air traffic control center. We recognize the importance of this observation by noting that in terminal area operation under instrument flight rules the control of the aircraft's position in space is often a cooperative effort between pilot and controller.

Some specific techniques for achieving the broad objectives of greater flight safety and operational efficiency are automating routine pilot and controller work in terminal area guidance, the introduction of advanced display concepts and avionics systems, and the use of a data link to connect computers in the aircraft and in the air traffic control center. The data link would greatly reduce the delays and inefficiencies inherent in the traditional verbal communications between the pilot and ATC. A specific function of such a data link will become evident shortly.

GUIDANCE PROBLEMS IN THE TERMINAL AREA

After this brief overview of a pilot-managed flight control system, a specific problem in its design is discussed. This problem deals with the automatic guidance of the aircraft in the terminal area. To illustrate this problem, Figure 1 shows, in

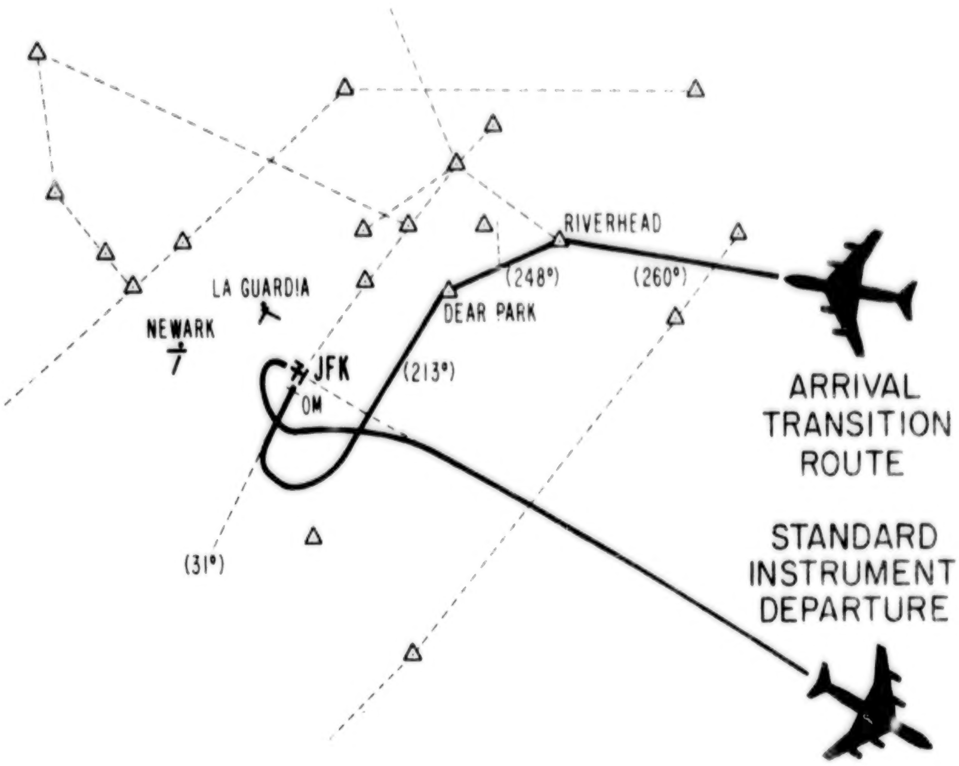


Figure 1.- Typical instrument departure and arrival routes (JFK)

simplified form, the air route structure and the instrument approach and departure procedures for oceanic operations at JFK International Airport. The dashed light lines together with the small triangles, denoting the location of VOR navigational stations, form a portion of the airway system in the neighborhood of the airport. An overseas aircraft on instrument approach flies toward the Riverhead VOR on a heading of 260°. Upon arrival there, it turns immediately toward the Deer Park VOR. At Deer Park it turns to a heading of 211° and after a period of straight line flight is directed by a controller observing the aircraft on radar to make a sequence of turns which brings it in alinement with the ILS localizer beam. The controller ensures that the aircraft is alined with the centerline of the beam before it reaches the outer marker and that adequate separation with other traffic is maintained. After ILS capture has occurred the pilot continues the approach on his own using on-board instruments.

Analysis of terminal arrival routes such as these suggests that the lateral guidance of the aircraft in the terminal area can be thought of as the solution of a sequence of three relatively simple guidance problems. Figure 2 defines these three problems in order of decreasing complexity. Problem (a) is defined as the

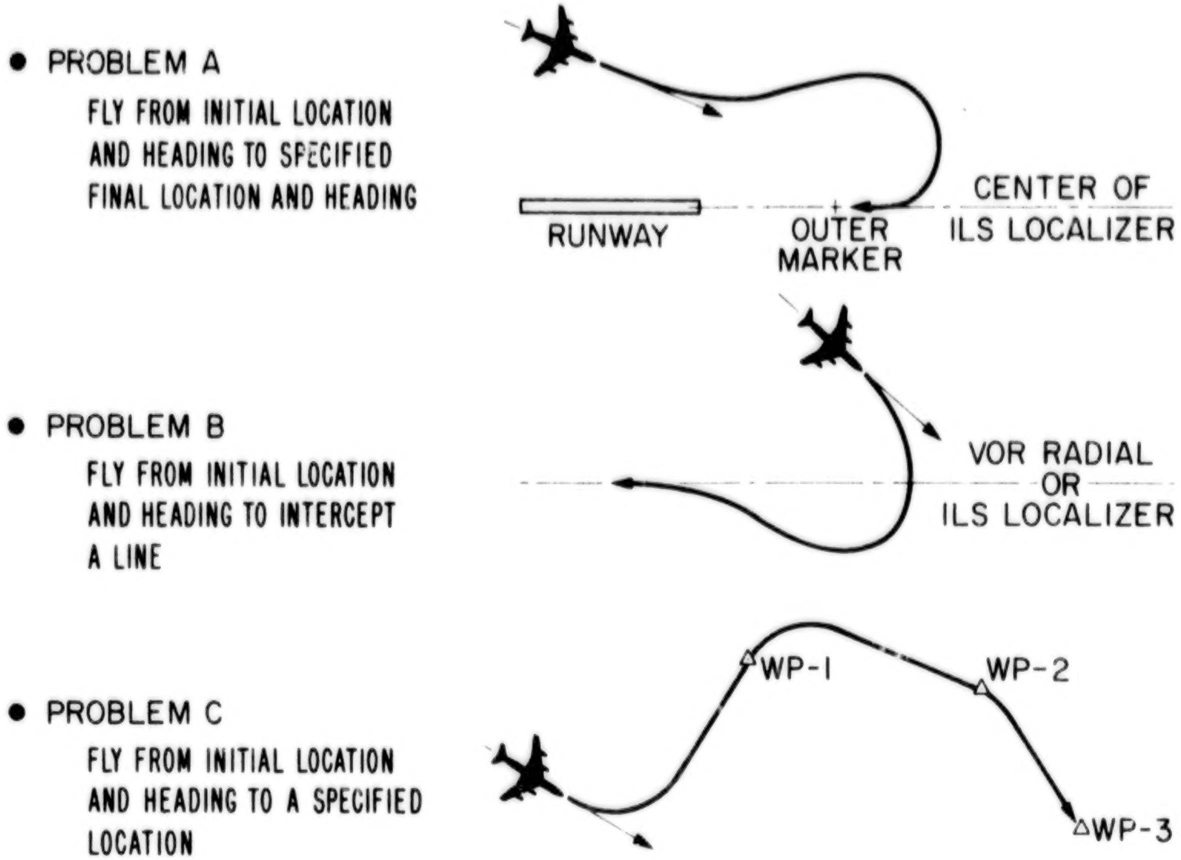

• PROBLEM A
FLY FROM INITIAL LOCATION AND HEADING TO SPECIFIED FINAL LOCATION AND HEADING
RUNWAY
OUTER MARKER
CENTER OF ILS LOCALIZER
• PROBLEM B
FLY FROM INITIAL LOCATION AND HEADING TO INTERCEPT A LINE
VOR RADIAL OR ILS LOCALIZER
• PROBLEM C
FLY FROM INITIAL LOCATION AND HEADING TO A SPECIFIED LOCATION
WP-1
WP-2
WP-3

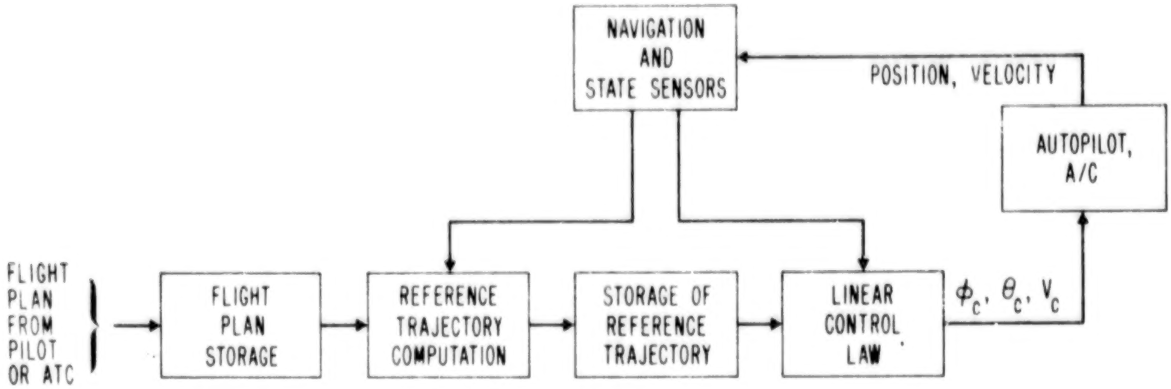
Figure 2.- Lateral guidance problems in the terminal area

guidance of the aircraft from an initial location and initial heading to a specified final location and final heading. It is the problem of acquisitioning the ILS localizer at the outer marker or some other specified point on the extended runway centerline. Problem (b) is defined as the guidance of the aircraft from an initial location and initial heading to intercept and fly along a directed line fixed in space. This problem is equivalent to capturing the localizer or a VOR radial without specifying the point of capture. Finally, problem (c) is defined as the guidance of the aircraft from an initial location and initial heading through a sequence of points called waypoints, where the heading of the aircraft over each point is not specified.

For practical purposes all guidance maneuvers occurring in the terminal area under routine operations can be synthesized as a sequence of these three problems. An analysis of vertical maneuvers in the terminal area is found to yield a corresponding set of problems. The vertical guidance problem is not discussed here since in most terminal area operations vertical guidance can be performed independently of horizontal guidance.

TECHNIQUE FOR TERMINAL AREA GUIDANCE

Figure 3 outlines the technique for on-board automatic terminal area guidance now being investigated at Ames. As an aircraft approaches the terminal area, air traffic control transmits a flight plan specifying the terminal approach route by means of a


NAVIGATION AND STATE SENSORS
POSITION, VELOCITY
AUTOPILOT, A/C
FLIGHT PLAN FROM PILOT OR ATC
FLIGHT PLAN STORAGE
REFERENCE TRAJECTORY COMPUTATION
STORAGE OF REFERENCE TRAJECTORY
LINEAR CONTROL LAW
ϕ_c, θ_c, V_c

CONSTRAINTS FOR LATERAL MANEUVERS:
- MINIMUM TURNING RADIUS

PERFORMANCE INDEX:
- ARC DISTANCE

Figure 3.- On-board terminal area trajectory computation and guidance

data link directly to the on-board guidance computer. Alternatively, the pilot may enter the flight plan manually. The flight plan, in line with the previous discussion, would be defined by a sequence of waypoints, enroute altitudes and arrival times at each waypoint. It is, of course, the responsibility of the air traffic control center to select the flight plan so it will not conflict with the operation of other air traffic in the area. The flight plan for the approach route stored in the on-board computer is then used together with the current position and velocity supplied by a navigation system to compute a reference trajectory which is also stored. A linear controller is used to maintain the aircraft's trajectory along the reference trajectory by sending appropriate roll, pitch, and speed commands to the autopilot.

COMPUTATION OF REFERENCE TRAJECTORIES FOR PROBLEM (a)

The computation of the reference trajectories between two adjacent waypoints is now illustrated for problem (a). In the horizontal plane, the constraint on the trajectories was taken as a specified lower limit on the minimum turning radius. At constant airspeed, limiting the minimum turning radius is equivalent to limiting the maximum bank angle and therefore the maximum normal acceleration. The performance index to be minimized in selecting the trajectories was chosen as the arc length.

The characteristics of the optimum controls and trajectories were established with the help of Pontryagin's Maximum Principle (ref. 1).* The optimum controls were found to be of three types, flying in a straight line, turning left with minimum turning radius, and turning right with minimum turning radius. Flying in a straight line corresponds to a singular arc in optimum control theory (ref. 2).

Solution for the optimum switching function yielded the totality of geometric patterns in the horizontal plane that can form optimum trajectories. Figure 4 lists these patterns for problem (a). The patterns on the left half of the figure contain turns and a straight line segment, those on the right turns only. It is seen that an optimum pattern may contain one straight line segment and if it does, there may be at most one turn at each end of the segment. These patterns always occur for distances of separation between initial and final points which are larger than four turning radii, though they may also occur for arbitrarily short distances. In patterns containing turns only, no more than four turns may occur. The interior turns of the two possible three-turn patterns must contain an angle larger than π radians, whereas the angles in the two interior turns of the four-turn pattern must be

---

*See appendix for equations of motion.

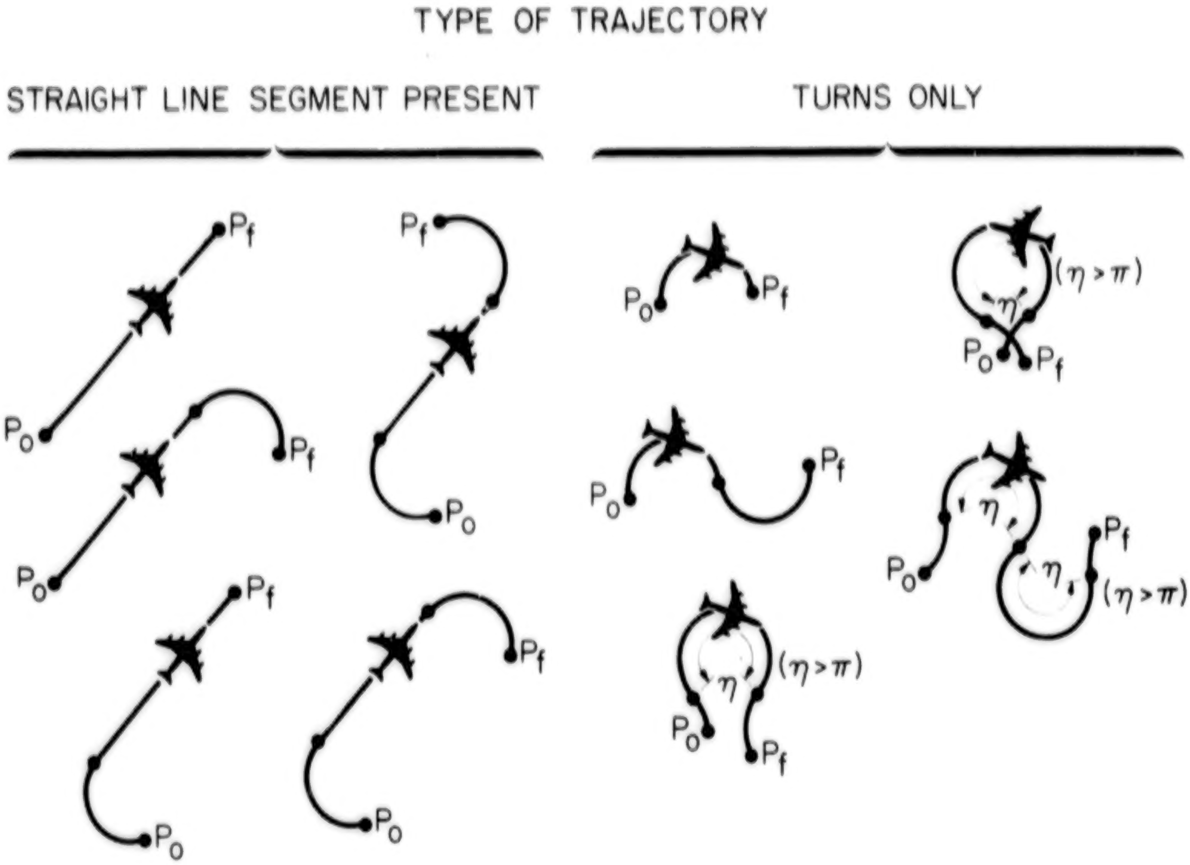


Figure 4.- Optimum trajectory patterns for problem (a)

equal and must also be larger than π radians. These "turns only" patterns occur only if the separation distance is equal to or less than four turning radii.

Specification of an optimum pattern for problem (a) requires knowledge of three states or parameters. These can be chosen in a number of different ways. We found that a convenient choice for computing purposes is to use the length of the directed line segment from the initial location to the final location as one parameter, and the initial heading measured with respect to the final heading as the third parameter.

Pilots examining the optimum trajectory patterns have expressed some concern with the operational acceptability of the three- and four-turn trajectories. In response to this concern a technique was developed that selects a suboptimum pattern consisting of no more than two turns and one straight line segment for any combination of the three parameters. This suboptimum technique has been implemented in a computer program.

OPTIMUM CONTROL LAW FOR PROBLEM (b)

For problem (b), defined earlier as flying to acquisition a directed line in the plane, a similar though less numerous set of optimum trajectory patterns is obtained. It is found that for this problem optimum patterns consist of no more than two turns and one straight line segment. These patterns can be used to synthesize the optimum control law which is illustrated in Figure 5.

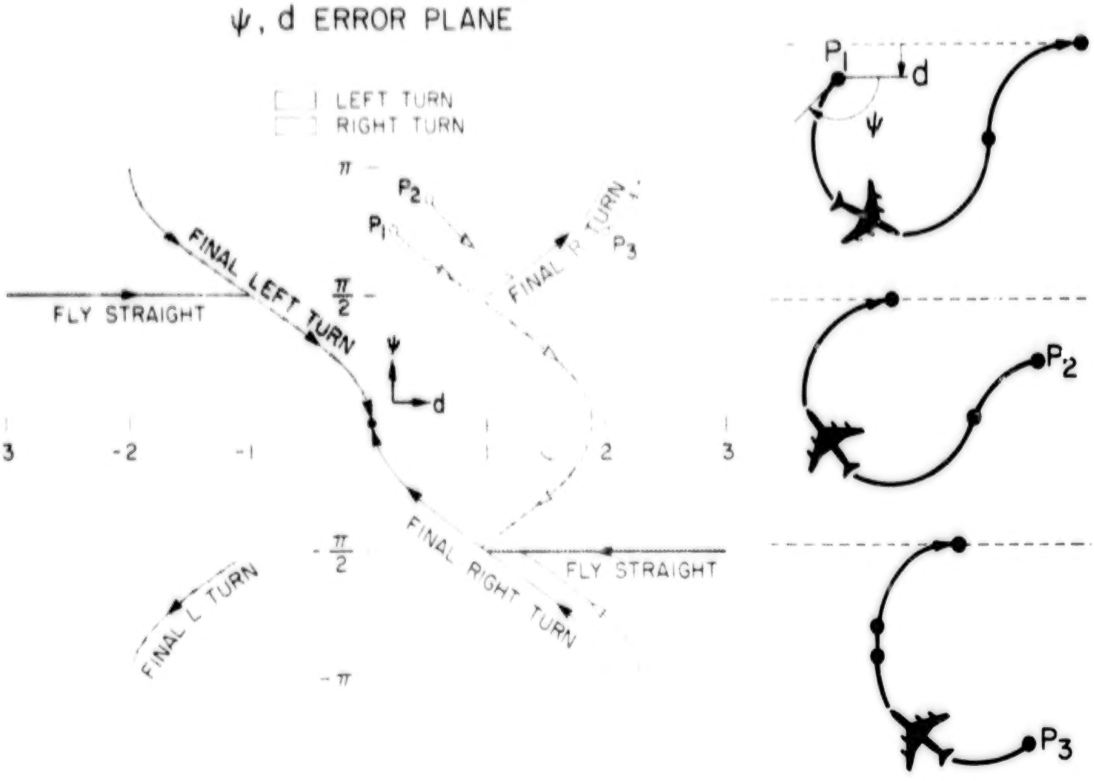

ψ, d ERROR PLANE
LEFT TURN
RIGHT TURN
π
P2
P1
FINAL R TURN
P3
FINAL LEFT TURN
π/2
FLY STRAIGHT
ψ
d
-3
-2
-1
1
2
3
FINAL RIGHT TURN
-π/2
FLY STRAIGHT
FINAL L TURN
-π
P1
d
ψ
P2
P3

Figure 5.- Control law for problem (b)

The control law is given here in terms of two parameters d and ψ, which are, respectively, the perpendicular distance of the aircraft's current position from the final line and the current heading angle of the aircraft measured with respect to the direction of the final line. Positive distances are associated with those aircraft positions lying to the right of the final line when facing in the direction of the line's heading. Positive heading angles are defined in the conventional manner.

The d, ψ parameter plane has been divided into three regions labeled left turn, right turn, and fly straight. The label of a region determines the control to be used for any parameter pair d and ψ falling in that region. The regions labeled right turn and left turn are bounded by sections of straight lines and portions of sine functions. The region labeled fly straight consists of two half lines which also form the trajectories in the parameter plane for this choice of control. The final turns onto the line and straight line flight are indicated here by solid lines, with arrows giving the direction of motion along the lines. Except for a segment of the final left turn between $\pi/2$ and π radians and a segment of the final right turn between $-\pi/2$ and $-\pi$ radians, these lines form a part of the boundary between the regions. The shading of the regions near the boundary formed by these lines has been exaggerated to give an indication into which region the line itself belongs. The part of the boundary not formed by these lines may be considered in either the left turn or the right turn region.

Figure 5 also shows the parameter trajectories and the corresponding trajectory patterns for three initial conditions. The trajectory starting at P_1 begins with a long left turn, and then moves to the origin with a final right turn. Starting at P_2, the trajectory makes a short left turn and concludes with a longer right turn. Although P_1 and P_2 lie in the same region and the trajectories start with the same control action, the complete trajectories are seen to differ substantially. Finally, the trajectory starting at P_3 begins with a long right turn, then flies straight for a short distance and concludes with a $\pi/2$ radian right turn.

CONCLUDING REMARKS

The proposed terminal area guidance technique can be summarized as follows. The air traffic control center transmits a conflict free-flight plan for terminal area operations to the arriving or departing aircraft. An onboard guidance computer synthesizes the precise trajectory from this flight plan according to agreed upon rules and guides the aircraft along it.

Three types of lateral guidance problems have been identified in synthesizing the terminal area trajectory in the horizontal plane.

Problem (a) was defined as the guidance of the aircraft from an initial location and heading to a specified location and specified final heading. Minimum arc distance trajectories for this problem contain no more than four minimum radius turns or no more than two turns and one straight line segment.

Problem (b) was defined as the guidance of the aircraft from an initial location and heading to intercept and fly along a directed line fixed in space. Minimum arc trajectories for this problem contain no more than two turns and one straight line segment.

Problem (c) was defined as the guidance of the aircraft from an initial location and heading through a set of waypoints, with the heading of the aircraft at each waypoint arbitrary. Minimum arc trajectories to fly between waypoints, though not discussed here in detail, contain no more than one turn and one straight line segment.

A computer simulation of the proposed terminal area guidance technique is currently being developed.

REFERENCES

1. Athans, M. and P. L. Falb: Optimal Control: An Introduction to the Theory and Its Applications, McGraw Hill, 1966, Ch. 5 & 6.

2. Johnson, C. D.: "Singular Solutions in Problems of Optimal Control." Vol. 2 of Advances in Control Systems, C. T. Leondes, ed., Academic Press, New York, 1965, pp. 209-267.

APPENDIX

EQUATIONS OF MOTION FOR HORIZONTAL GUIDANCE

A simplified set of equations is used in this study to describe the motion of the aircraft in the horizontal plane. It is assumed that the aircraft flies at small flight-path angles during the horizontal maneuvers and that the turns achieved by banking the aircraft are coordinated with the yaw rate to maintain zero sideslip. The roll time constant is neglected and thrust is set equal to drag. Under these conditions the equations of motion in the horizontal plane are as follows:

$$\frac{dx}{dt} = v_x$$

$$\frac{dy}{dt} = v_y$$

$$\frac{dv_x}{dt} = \frac{u}{m} \sin \psi$$

$$\frac{dv_y}{dt} = -\frac{u}{m} \cos \psi$$

In these equations, m is the aircraft mass, $\psi \equiv \tan^{-1} v_y/v_x$, and u is the horizontal force which is related to the bank angle ϕ by $u = w \tan \phi$, where w is the aircraft weight. The bank angle ϕ is taken as the control variable. For passenger comfort its magnitude is required to be less than ϕ_{max}.

The trajectories given in this study minimize the arc length of the trajectories between specified initial and final conditions. It can be shown that minimizing the arc length is equivalent to minimizing the transition time with the additional restriction that the speed be constant, $\sqrt{v_x^2 + v_y^2} = v$, where v is a constant of the motion. The minimum turning radius constraint can then be converted into a maximum bank angle constraint by the relation

$$R_{min} = \frac{v^2}{g \tan \phi_{max}}$$

Application of the Maximum Principle to this equivalent minimum time problem yields the results described in the paper.

INERTIALLY-AIDED, ALL-WEATHER LANDING SYSTEMS FOR C/TOL

By Robert J. Pawlak
Electronics Research Center

DISCUSSION

An attempt to find ways to improve significantly all-weather landing systems technology for conventional take-off and landing vehicles, primarily, the large commercial jets, has been made at the Electronics Research Center for the last two and one-half years. The majority of these vehicles have well-designed flight control systems and some degree of automatic approach and landing capability. Many of them also have, or will have in the future, an inertial navigation system on-board. Presently, these navigators are used only for navigation and are generally thought to be of no use for the landing problem which demands a few feet of navigation precision as compared to what is typically thought to be the one or two or many nautical miles of position error of these navigators at the end of a typical flight. After this lengthy study, it has been concluded that there is a great deal to be gained by extensive use of inertial navigation system derived information, primarily, the translational information inherent in an inertial navigator, for improving all-weather landing system performance.

An analysis and simulation effort has been underway at ERC to develop such concepts for about the last two years. In April 1969, analysis results were significant enough to justify a flight test program to verify some of the basic concepts evolved through the analysis and simulation effort.

At present, NASA ERC has contracted to design, implement, install, and flight test a "minimum modifications" Phase-I-type system in the FAA's Convair 880, a large four-engine jet operating out of NAFEC, the FAA's flight test range near Atlantic City, New Jersey. This is basically a joint NASA-FAA Program using NASA research and development dollars and the FAA's aircraft, flight crews, facilities, and services at the NAFEC test range. The minimum modifications design approach corresponds to the sort of system that could be retrofitted in existing vehicles with the addition, essentially, of only an interface box between the inertial navigation system and the existing approach and landing system. The design goal in this effort is to achieve, at least, a-factor-of-2 increase in dynamic performance using the new techniques; that is, the improved ability to track the ideal trajectory for minimum dispersions at all altitudes down to touch down. Flight test results are expected between February and May 1970.

Meanwhile, the analysis and simulation effort has changed from the SST model used up until last April to the Convair 880 vehicle model dynamics. Phase II design, which is more consistent with the previous analysis and simulation work, is presently underway. This work assumes the availability of a good-size, on-board computer, in addition to the simple interface that we normally associate with the Phase I effort. Phase III studies, which will develop means for using the basic Phase II system capability, are also underway. These developments will allow for such things as path and time-constrained terminal area trajectories, decelerating approaches for noise abatement, curved approaches to ILS beam interception for maximum air field acceptance rates, improved automatic decrab, deceleration after touchdown, roll out, high speed turn off, and other pertinent considerations.

The Phase I design for the Convair 880 will not be available until the first week in December. By February or March of next year, however, a large matrix of simulation results will be available. It will compare the effects of different possible combinations of navigators of different qualities, ILS beams of different qualities, wind disturbances of different intensities, and different levels of control system design optimization, all with the same data base of the 880 as the vehicle model. It is hoped that somewhere in this large matrix of analysis results will reveal flight test verification of that small subset of conditions corresponding to the particular Phase I design implemented in the Convair 880. This kind of flight test verification of analytical results will give a great deal of confidence in the total matrix of results.

The following figures correspond to just the lateral channel, Phase II-type design concepts. Phase II requires an on-board computer as opposed to Phase I which only uses an interface box. A definition of what these concepts are, and what one would expect to achieve in the way of system improvement by using them, will be presented as part of the explanation of each figure.

Figure 1 lists the typical aircraft flight control system sensors used to build the pitch, roll, yaw, attitude, and throttle control systems of a typical aircraft. Down the left side of the figure are the sensors, grouped according to the particular control subsystem in which they are used. Except for altimeters, essentially no translational state information is gathered with the sensors. That is the important message of this particular figure. Even the ILS localizer receiver and the ILS glideslope receiver information, as derived on-board the aircraft, is also angular information. It can only be considered translational information if one knows the precise range to the glideslope transmitter or localizer transmitter. Therefore, typical aircraft flight control systems are primarily using angular state information.

Figure 2 presents a typical gimbaled inertial navigation system which, when thought of as a sensor, generates lots of very good translational state information. It may not be correct information in an absolute sense, but the relative position and relative velocity information for short periods of time, such as 5 or 10 min during a landing phase, is excellent. It should be noted that the aircraft flight control system's body-mounted rate gyros are usually used in preference to inertial system angular rate information, but for absolute attitude and heading, the inertial navigator is far superior to anything else on board. Significantly, the figure shows that if the control system designer is given inertial navigation system-derived translational state information, in addition to the angular state information, he should be able to come up with significantly better approach and landing systems.

Figure 3 shows the NASA-ERC overall scheme of things for inertially-aided, all-weather landing systems. The basic philosophy is to command simultaneously all three derivatives of the vehicle's translational state, fly relative to the best on-board estimate of the translational state, generate the best estimate of that state by blending inertial navigation system-derived data with VOR, DME, ATC radar fixes, or any other means in the general enroute and near terminal area, and then blend ILS information with the inertial data for this best estimate during final approach and landing. At the bottom of the figure the command is shown to be coming out of the non-linear trajectory generator, described later. This command consists of simultaneous position, velocity, and acceleration state information. The best estimate of the vehicle's state is generated in the estimator and comes out as all three derivatives of x, y, and z shown as $\overline{X}$, $\overline{Y}$, and $\overline{Z}$. The angular information that is processed around the loop is primarily the angular information that already exists in the typical flight control system. Given the best estimate of on-board state and the desired vehicle state, the flight control system designed then generates the best command of the effectors and the throttle shown as δ_E and δ_T.

This figure shows that there are three primary disturbances with which one must contend. The first is aerodynamic noise acting on the vehicle, i.e., wind, wind gust, and wind gradient disturbances; ground effects; and the like. The second is different kinds of sensor noise that acts on all the various sensors being used to help generate the best estimate of vehicle state. The third set of disturbances are the inertial navigation system error sources. An arrow called calibration and compensation is shown coming back from the estimator. Other work is proceeding at NASA-ERC to help up-date the inertial system in the true sense at some later date. That kind of information is not being used in the CTOL Program yet, but estimators of the type shown in Figure 4 are being used.

PITCH ATTITUDE GYRO	ELEVATOR DEFLECTION (PITCH CS) (ALTITUDE CS)
PITCH RATE GYRO	
BAROMETRIC ALTIMETER	
RADAR ALTIMETER	
ILS GLIDESLOPE RECEIVER	
NORMAL ACCELEROMETER	
($\dot{H}$ SENSOR)	
ROLL ATTITUDE GYRO	AILERON DEFLECTION (ROLL CS)
ROLL RATE GYRO	
HEADING	
ILS LOCALIZER RECEIVER	
YAW ATTITUDE GYRO	RUDDER DEFLECTION (YAW CS)
YAW RATE GYRO	
HEADING	
ILS LOCALIZER RECEIVER	
PITCH ATTITUDE GYRO	THRUST (THROTTLE CS)
AIR SPEED SENSOR	

Figure 1.- Typical aircraft flight control system

$\ddot{X}$ $\ddot{Y}$ $\ddot{Z}$	LAGGED BY $\tau = 0.1$ SEC
$\dot{X}$ $\dot{Y}$ $\dot{Z}$	EXCELLENT
X Y Z	EXCELLENT
$\ddot{\theta}$ $\ddot{\phi}$ $\ddot{\psi}$	NOT AVAILABLE
$\dot{\theta}$ $\dot{\phi}$ $\dot{\psi}$	AFCS RATE GYROS BETTER
θ ϕ ψ	BEST ATTITUDE & HEADING REFERENCE

Figure 2.- Typical gimbaled inertial navigation system sensor

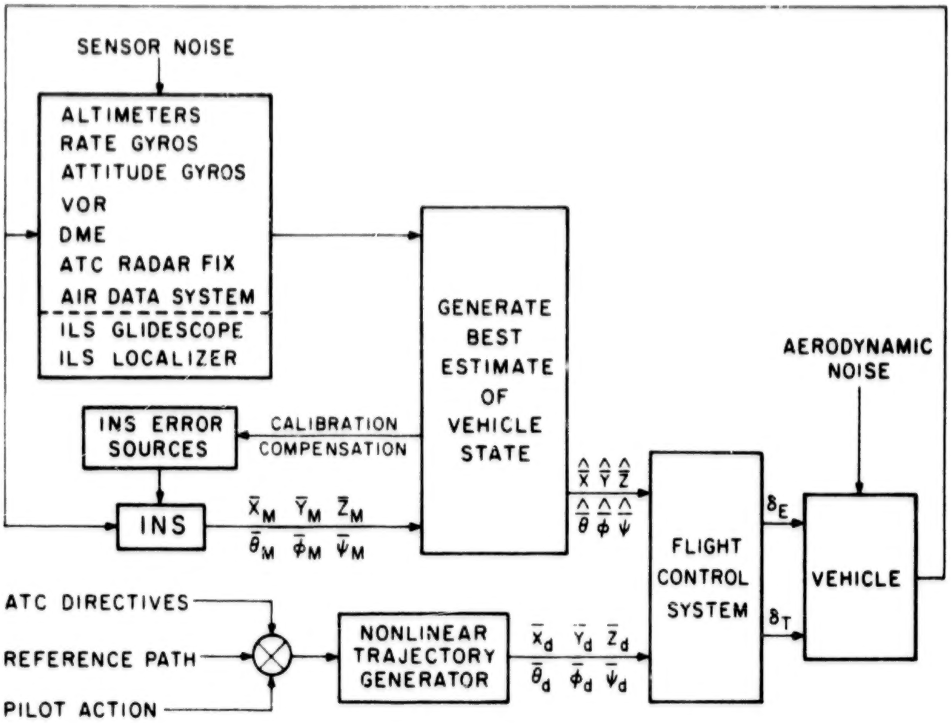

SENSOR NOISE
ALTIMETERS
RATE GYROS
ATTITUDE GYROS
VOR
DME
ATC RADAR FIX
AIR DATA SYSTEM
ILS GLIDESCOPE
ILS LOCALIZER
GENERATE BEST ESTIMATE OF VEHICLE STATE
INS ERROR SOURCES
CALIBRATION COMPENSATION
INS
$\bar{X}_M$ $\bar{Y}_M$ $\bar{Z}_M$
$\bar{\theta}_M$ $\bar{\phi}_M$ $\bar{\psi}_M$
$\hat{X}$ $\hat{Y}$ $\hat{Z}$
$\hat{\theta}$ $\hat{\phi}$ $\hat{\psi}$
AERODYNAMIC NOISE
FLIGHT CONTROL SYSTEM
δ_E
δ_T
VEHICLE
ATC DIRECTIVES
REFERENCE PATH
PILOT ACTION
NONLINEAR TRAJECTORY GENERATOR
$\bar{X}_d$ $\bar{Y}_d$ $\bar{Z}_d$
$\bar{\theta}_d$ $\bar{\phi}_d$ $\bar{\psi}_d$

Figure 3.- Inertial navigation system aided all-weather landing system

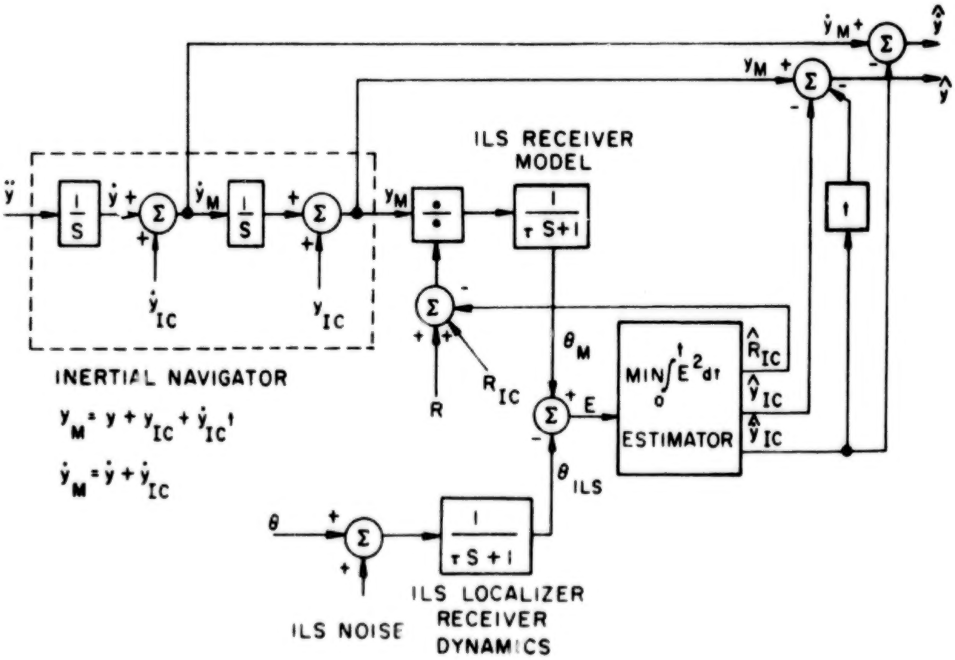

$\ddot{y}$
$\frac{1}{S}$
$\dot{y}$
$\dot{y}_M$
y_M
$\dot{y}_{IC}$
y_{IC}
INERTIAL NAVIGATOR
$y_M = y + y_{IC} + \dot{y}_{IC}t$
$\dot{y}_M = \dot{y} + \dot{y}_{IC}$
ILS RECEIVER MODEL
$\frac{1}{\tau S+1}$
R
R_{IC}
θ_M
E
θ_{ILS}
$MIN\int_0^t E^2 dt$
ESTIMATOR
$\hat{R}_{IC}$
$\hat{y}_{IC}$
$\hat{\dot{y}}_{IC}$
t
$\hat{y}$
$\hat{\dot{y}}$
θ
ILS NOISE
ILS LOCALIZER RECEIVER DYNAMICS

Figure 4.- Lateral state estimator

Figure 4 shows an overall block diagram of the lateral state estimator, in which Y is the lateral displacement, the distance from the ideal localizer plane that contains the runway center line; and θ_{ILS}, as shown at the bottom of the figure, is the ILS localizer receiver output. It is the measurement of the angle between the localizer plane and the ray from the localizer transmitter to the aircraft. It should be noted that it is degraded by ILS noise, and that a lag is used as the model of the ILS localizer receiver dynamics. θ_M is an inertial navigation system-derived equivalent measurement as deteriorated primarily by three errors. The inertial navigator is modelled in a simple fashion as shown by the two equations on the figure. This is an effective model for short-term behavior of the navigator. Starting with an initial condition error in velocity and an initial condition error in position, one must divide by an estimate of range to the localizer receiver, as shown by the sum of R and R_{IC}, to get an equivalent angular measurement. R is the true range and R_{IC} is the inability to perfectly estimate R initially. To get an exactly equivalent angular measurement, the inertial navigation system measurement channel is degraded by the same dynamics that are in the ILS channel, namely, an ILS receiver model. An error signal E that is the difference between these two measurements is formed. Hopefully, they are measuring the same thing, and this error signal passes into the estimator. A minimum integral error squared estimator has been shown. This type of estimator has been mechanized, as has a Kalman filter. The result is some estimate of $\hat{R}_{IC}$, $\hat{y}_{IC}$, and $\hat{\dot{y}}_{IC}$. These signals, when properly combined with y_M and $\dot{y}_M$, lead to the best estimate of velocity and best estimate of position as shown in the upper right-hand corner of the figure as $\hat{y}$ and $\hat{\dot{y}}$. The primary reason why this whole procedure works well is the fact that the estimator process is fast relative to the rate at which y_{IC}, $\hat{\dot{y}}_{IC}$ change as the ideal characterization of an actual navigator.

The non-linear trajectory generator shown in Figure 5 is for the lateral channel and is simply a critically-damp second-order model whose dynamics are slower than the gross vehicle dynamics. The limiters are selected such that the actual response of the vehicle to y_d, $\dot{y}_d$, and $\ddot{y}_d$ commands that come out of the non-linear trajectory generator will not saturate the angle limits and angular rate limits of the vehicle's effectors. This leaves enough effector travel to handle the higher frequency disturbance rejection commands that are superimposed on the lower frequency gross trajectory commands shown here. The effect of the non-linear trajectory generator is shown later in Figures 7 and 8.

Figure 6 shows one particular lateral position control system on which NASA-ERC has worked. The input, y_c goes into the non-linear trajectory generator. y_d, $\dot{y}_d$, and $\ddot{y}_d$ are the desired

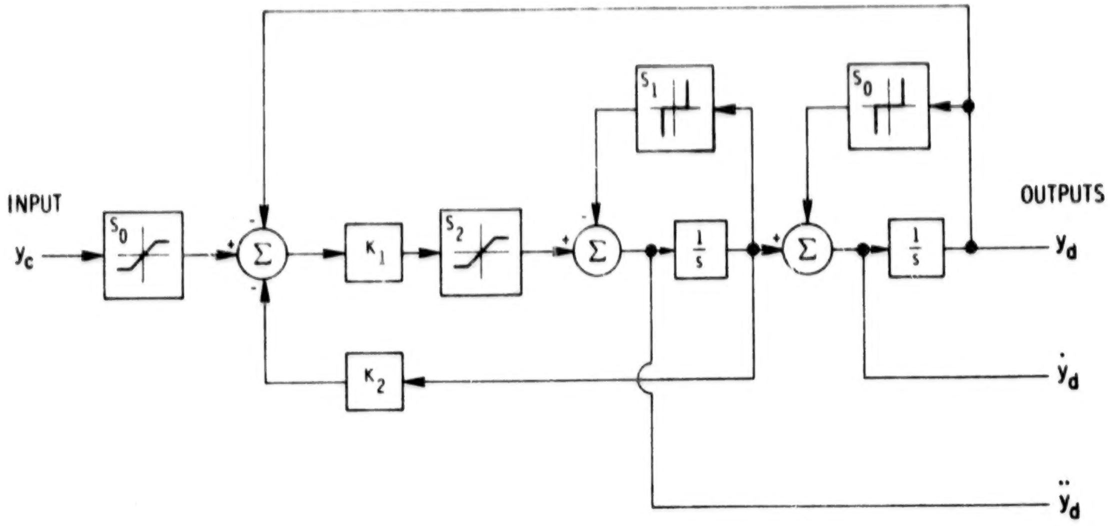
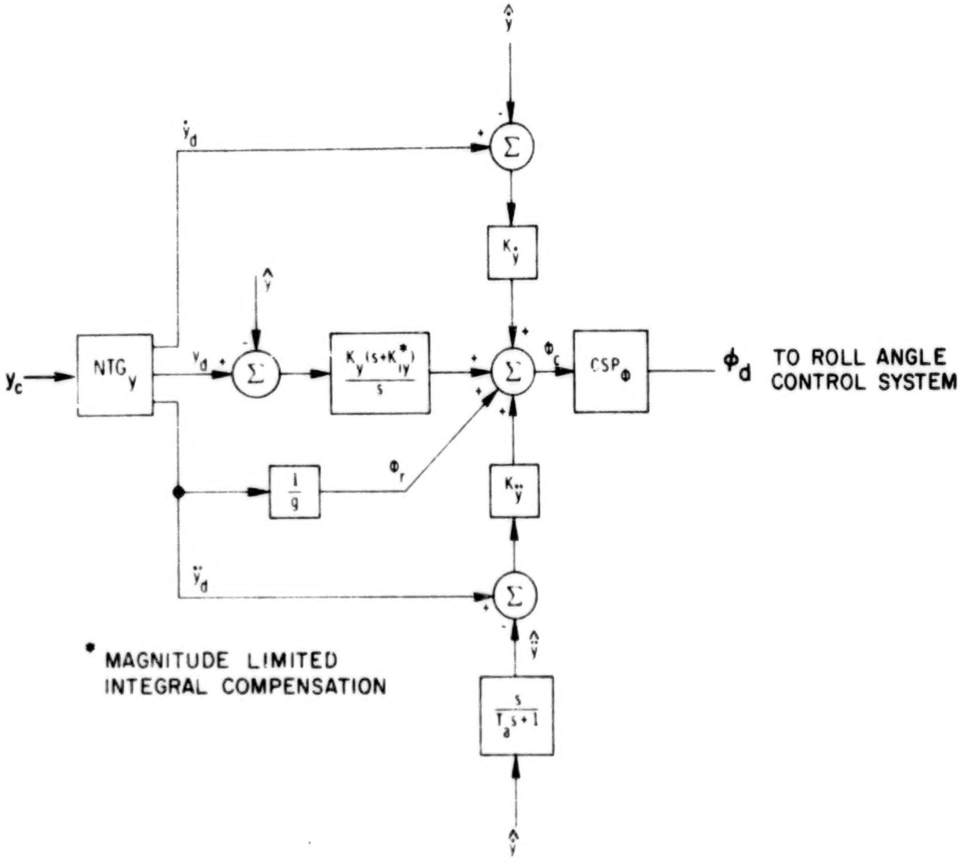

INPUT
y_c
S_0
K_1
S_2
S_1
S_0
$\frac{1}{s}$
$\frac{1}{s}$
K_2
OUTPUTS
y_d
$\dot{y}_d$
$\ddot{y}_d$

Figure 5.- Non-linear trajectory generator


$\hat{\dot{y}}$
$\dot{y}_d$
$K_{\dot{y}}$
$\hat{y}$
y_c
NTG_y
y_d
$\frac{K_y(s+K^*_{iy})}{s}$
ϕ_c
CSP_ϕ
ϕ_d
TO ROLL ANGLE CONTROL SYSTEM
$\frac{1}{g}$
ϕ_r
$K_{\ddot{y}}$
$\ddot{y}_d$
$\hat{\ddot{y}}$
$\frac{s}{T_a s+1}$
$\hat{\dot{y}}$

* MAGNITUDE LIMITED INTEGRAL COMPENSATION

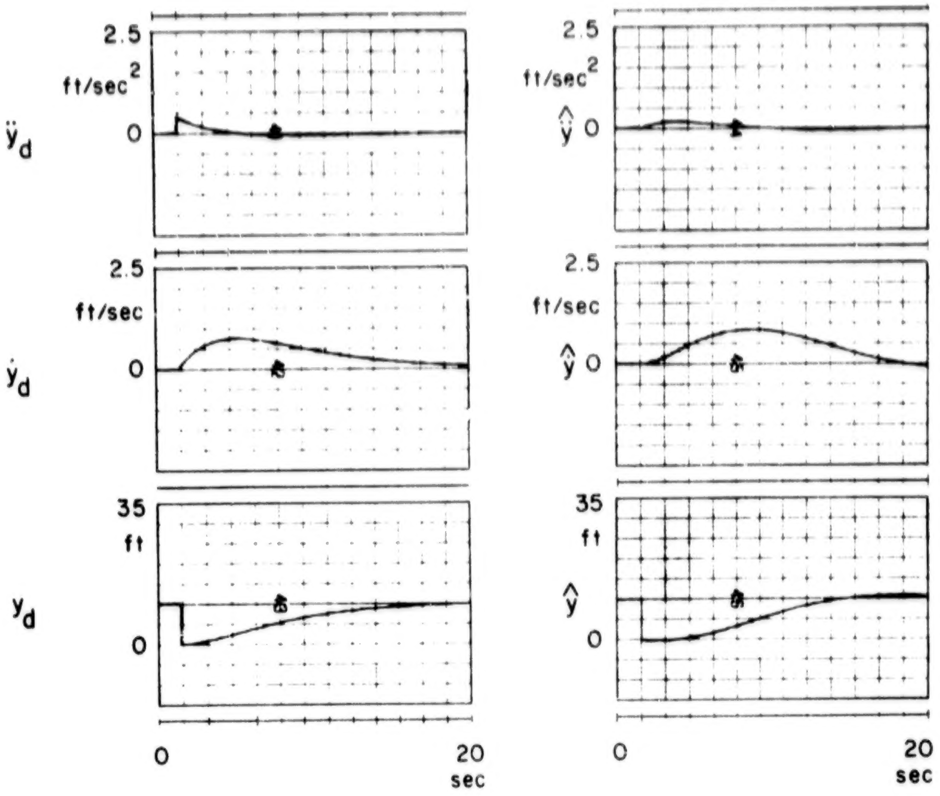
Figure 6.- Lateral position control system

values of translational state coming out of the non-linear trajectory generator. $\hat{y}$ and $\hat{\dot{y}}$ are the best estimate of lateral state coming out of the estimator as shown in Figure 4. At the bottom of the figure, $\hat{\dot{y}}$ goes through a transfer function which differentiates the signal at the expense of a lag which has a time constant of about 0.1 sec, giving a best estimate of translational acceleration Now the control system designer has available the simultaneous command of the full translational state and the best estimate of on-board measured translation state. He can now operate on the difference between these two or the absolute of either one to form his control system design. The choice of the gains, compensations, and the like are his option, but this is quite a luxury of feedback information. The result is a roll command that goes to the roll angle control system, the basic interloop of the vehicle.

Figure 7 shows the lateral position control system response. y_d, $\dot{y}_d$, and $\ddot{y}_d$ are the outputs of the non-linear trajectory generator for a 10-ft step change in Y_c, the basic initial command into the system; y, $\hat{\dot{y}}$, and $\hat{\ddot{y}}$ are the responses of the lateral position control system, and show the systems ability to follow the simultaneous y_d, $\dot{y}_d$, and $\ddot{y}_d$ state commands. Obviously, the system follows the total state command quite well.

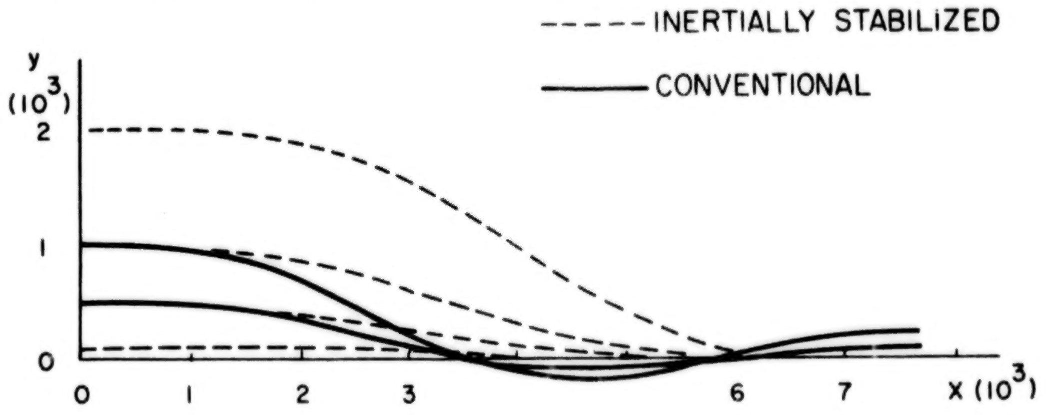
Figure 8 shows the comparison of the inertially stabilized lateral control system response with that of the conventional system response for y_c step changes. The inertially stabilized system has a slower response in a natural frequency sense, but there is no overshoot which is especially important when you have parallel runway approaches a few thousand feet apart. The responses are slower, but they do not have the oscillatory tendency of the conventional system. It is also important to note that the inertially-stabilized system is capable of eventually achieving large step commands without saturating any of the effector limits, whereas in the standard or conventional control system, as shown here, a 2000-ft step command, as the system was mechanized, would have come up against the roll angle of 30 deg, resulting in very large overshoots not shown here because it is a highly undesirable situation.

Basically, Figure 9 shows the response to ILS localizer beam noise, comparing the conventionally stabilized lateral position and the inertially stabilized lateral position and then the roll angles corresponding to the two types of systems. Furthermore, the figure basically says that the inertially stabilized system is much less susceptible to ILS beam noise, primarily, because the estimator is doing its job properly. As you would expect, the roll angle activity corresponding to the two systems shows a considerable reduction in roll angle activity for the inertially stabilized system. In the inertially stabilized case, the vehicle



10 FT STEP CHANGE IN y_c AT $t = 0$

Figure 7.- Lateral position control system response



NOTE: CONVENTIONAL RESPONSE TO 2000 FT COMMAND
NOT SHOWN SINCE 30° ROLL ANGLE LIMIT
WAS EXCEEDED

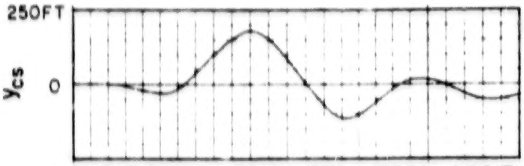
Figure 8.- Lateral control system response to step command in Y

is essentially flying straight and level against good guidance information whereas, in the conventional case, the aircraft is chasing the ILS beam noise all over the sky.

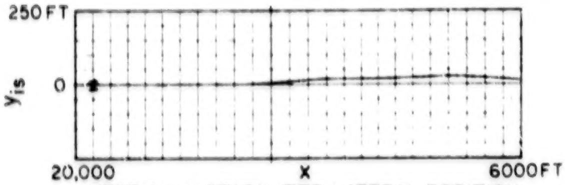
Figure 10 shows the lateral response to wind gust and linear wind gradients. This figure basically says that the inertially stabilized system is much less susceptible to wind gust and wind gradient disturbance. The wiggles represent the responses to gust. The gross tendency to drive the vehicle off the localizer plane or y = 0 line is the response to wind shear. For the conventionally stabilized case, the vehicle is being driven away from the localizer plane at a steady rate. Whereas, in the inertially stabilized case, the vehicle has been driven off as far as it is going to go and the integrator and by-pass in that channel and is about to bring the vehicle back to zero. In any case, for the conditions shown here where the vehicle has flown from a 1000 to 200 ft in altitude, the upper chart would be an unacceptable situation with the vehicle already 20 ft off the center line at 200 ft altitude and drifting further from the lateral plane as it descends.

Figure 11 just says that the roll angle activity corresponding to Figure 10 is again significantly reduced for the inertially-stabilized system. Most of the activity shown here is in response to wind gusts, not the wind gradient.

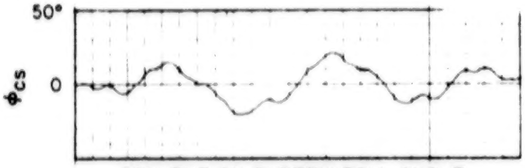
Figure 12 shows lateral touchdown dispersions as a function of mean squared turbulence intensity. The final proof of such a system here being developed is the degree of decrease in lateral touchdown dispersion that can be expected by an inertially stabilized system. This figure shows the considerable improvement one can expect as a function of just one of the variables, namely,turbulence intensity. It is this kind of result that leads us to the basic conclusion that we should be in a position to achieve at least a factor of about 4 in dynamic performance improvement with a Phase II type design system mechanization such as described in this paper, and perhaps a factor or two in performance improvement with a Phase I type design mechanization, which will be flight-tested in the 880.



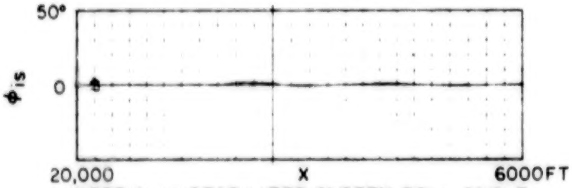
CONVENTIONALLY STABILIZED LATERAL POSITION



INERTIALLY STABILIZED LATERAL POSITION

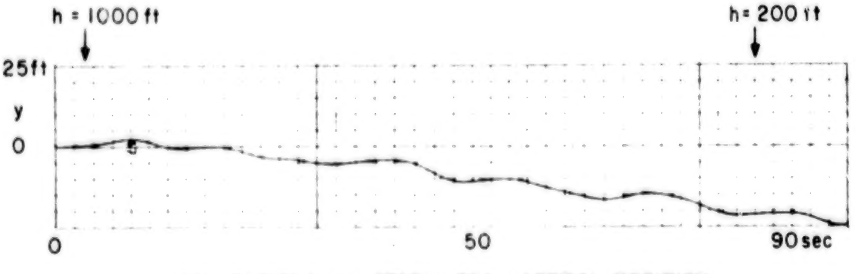


CONVENTION SYSTEM ROLL ANGLE

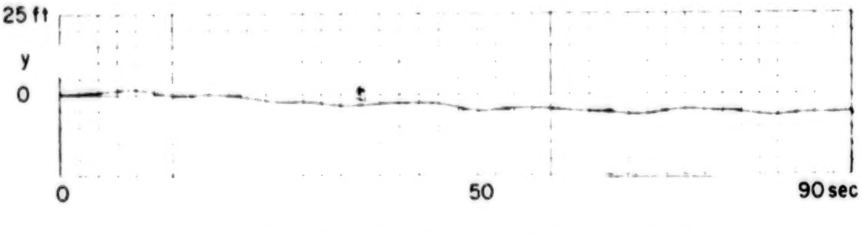


INERTIALLY STABILIZED SYSTEM ROLL ANGLE

Figure 9.- Response to instrument landing system localizer beam noise



CONVENTIONALLY STABILIZED LATERAL POSITION



INERTIALLY-STABILIZED LATERAL POSITION

Figure 10.- Lateral response to wind gusts and linear wind gradient

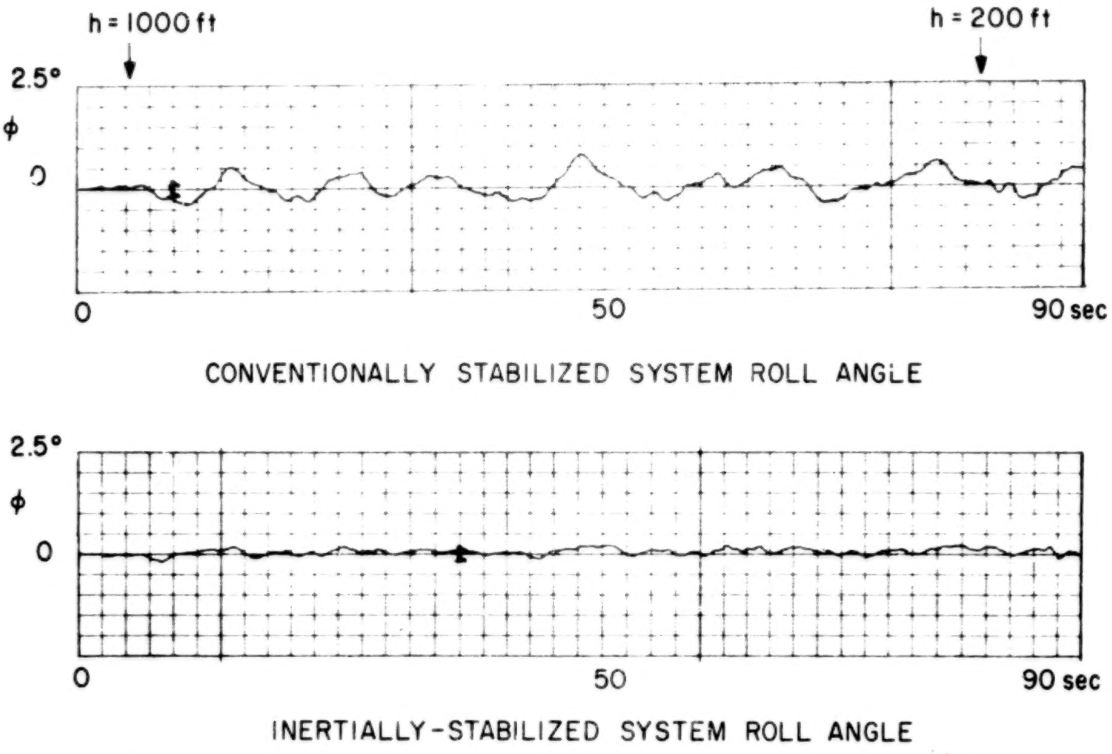


Figure 11.- Roll angle response to wind gusts and linear wind gradients

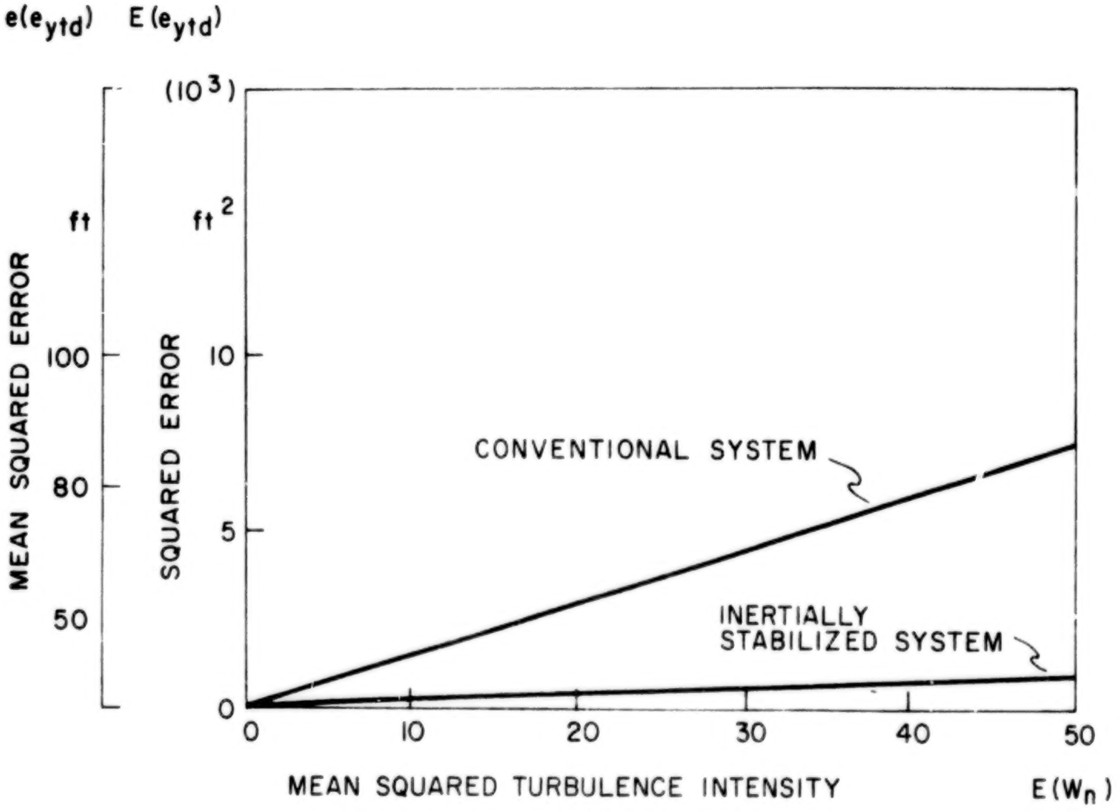


Figure 12.- Lateral touchdown dispersions

REFERENCES

Broxmeyer, C., MacKinnon, D., and Madden, P.: "Application of Inertial Navigation and Modern Control Theory to the All-Weather Landing Problem." Report R-613, M.I.T. Instrumentation Laboratory, Cambridge, Mass., June 1968.

MacKinnon, D.: "Improving Automatic Landing System Performance Using Modern Control Theory and Inertial Measurements." Report R-628, M.I.T. Instrumentation Laboratory, Cambridge, Mass., January 1969.

MacKinnon, D.: "Some Applications of Mathematical Optimization to Automatic Landing Systems." Report R-651, M.I.T. Instrumentation Laboratory, Cambridge, Mass., December 1969.

BLANK PAGE

AN APPROACH TO THE DETERMINATION OF AIRCRAFT HANDLING QUALITIES USING PILOT TRANSFER FUNCTIONS

James J. Adams and Howard G. Hatch, Jr.
Langley Research Center

SUMMARY

It is shown that a correlation exists between pilot-aircraft system closed-loop characteristics, determined by using analytical expressions for pilot response along with the analytical expression for the aircraft response, and pilot ratings obtained in many previous flight and simulation studies. Two different levels of preferred pilot response are used. These levels are:

1. A static gain and a second-order lag function with a lag time constant of 0.2 second;
2. A static gain, a lead time constant of 1 second, and a 0.2-second lag time constant.

If a system response with a pitch-angle time constant of 2.6 seconds and a stable oscillatory mode of motion with a period of 2.5 seconds can be achieved with the first-level pilot model, it is shown that the pilot rating will be satisfactory for that vehicle. Further, if an altitude response with a stable oscillatory mode of motion with a period of 5 seconds can be achieved, the vehicle will be rated satisfactory. If the second level pilot model is required to achieve these system response characteristics, the aircraft will be rated acceptable-unsatisfactory.

INTRODUCTION

It has been shown in references 1 and 2 that models of human response can be used to predict system performance of manually controlled vehicles. It would be beneficial if these pilot models could also be used to predict pilot acceptance of future vehicles. Since a large amount of previously obtained data on pilot ratings of aircraft as a function of aircraft characteristics is available, criteria for predicting pilot ratings are established in this paper by showing the correlation between calculated pilot-vehicle system characteristics, determined by using pilot models, and the previously obtained pilot ratings of aircraft. The criterion that is developed involves specified levels of preferred pilot response characteristics and specified pilot-vehicle closed-loop system characteristics. These criteria are suggested for use in predicting pilot ratings of future vehicles.

SYMBOLS

δ_e	control deflection, rad
ε	displayed error, rad
K_1	pilot model static gain
T_1	pilot model lead time constant, sec
T_2	pilot model lag time constant, sec
s	Laplace operator, per sec
θ	pitch angle, rad
x	displacement, ft (m)
h	altitude, ft (m)
T_{θ_2}	open-loop aircraft lead time constant, sec
M_α	normalized pitching-moment derivative, per sec
M_q	normalized damping-moment derivative in pitch, per sec
M_{δ_e}	normalized pitching control-moment derivative, per sec
L_α	normalized lift-force derivative, per sec
L_{δ_e}	normalized control-lift derivative, per sec
ω_n	open-loop aircraft undamped natural frequency, rad/sec
ω_f	control actuator undamped natural frequency, rad/sec
α	angle of attack, rad
V	velocity, ft/sec (m/sec)
ζ	damping ratio
T	closed-loop time constant, sec
P_h	closed-loop period of the altitude mode of motion, sec
P_α	closed-loop period of the angle-of-attack mode of motion, sec
ζ_h	damping ratio of the altitude mode of motion

ζ_α	damping ratio of the angle-of-attack mode of motion
γ	flight-path angle, rad

SUBSCRIPTS

ε	error
c	command
h	altitude
θ	pitch

PILOT MODELS

The pilot models which will be used in this paper are described in the following manner. For single-loop control tasks, the model form used consists of the following elements: a static gain and a lead time constant which constituted a command control deflection; and a second-order, critically damped lag function, which represents the dynamic response of the arm in executing the command for the control deflection. In transfer function form, the model is:

$$\delta/\varepsilon = \frac{K_1\left(1 + T_1 s\right)}{\left(1 + T_2 s\right)^2} \tag{1}$$

Typical values for the coefficients for various plant dynamics are given in Table I. These values are taken from references 3 and 4.

TABLE I. - MEASURED PILOT RESPONSE

Dynamics	Pilot transfer function		
$2/s$	$\frac{3.26}{(1 + 0.25s)^2}$	$\frac{3.2(1 + 0.33s)}{(1 + 0.33s)^2}$	$\frac{0.7(1 + 0.67s)}{(1 + 0.33s)^2}$
$\frac{10}{s(s + 1)}$	$\frac{0.5(1 + 0.3s)}{(1 + 0.1s)^2}$	$\frac{0.36(1 + 0.91s)}{(1 + 0.18s)^2}$	
$\frac{10}{s^2(s + 1)}$	$\frac{0.29(1 + 1.0s)}{(1 + 0.05s)^2}$		

It can be seen from this table that when a pilot is controlling a very easy-to-handle plant, such as a pure rated vehicle, 2/s, the lag time constant of the pilot's response, T_2, is as large as 0.3 second, with 0.2 second being a very common average value. Also, the lead time constant, T_1, is very often zero. However, sometimes a value of up to 0.67 second is measured. As the control difficulty of the plant is increased by increasing the lag of the plant, additional lead is added to the pilot's response, and the pilot's lag is reduced. In the extreme case of a plant $10/[s^2(s+1)]$, the lead time constant is increased to the maximum measured value of 1.0 second, and the lag time constant is reduced to 0.05 second.

In one experiment reported in reference 3, pilot response was measured while controlling a plant with an aircraft pitch response type of characteristic. The plant transfer function was

$$\frac{\theta}{\delta} = \frac{10}{s(s^2 + 3s + 10)} \tag{2}$$

and the measured pilot transfer function was

$$\frac{\delta}{\theta_c} = \frac{0.86(1 + 0.7s)}{(1 + 0.14s)^2} \tag{3}$$

It can be seen from this transfer function that the lag and lead time constants are within the ranges noted in the previous paragraph. It is assured that the values given for maximum and minimum time constants given before will also apply to aircraft type of plant dynamics.

The static gain, K, is always adjusted to provide a desired system response characteristic. That these system characteristics can be identified and defined is shown by the results obtained in reference 3. In reference 3, a wide variety of plants, control gains, display gains, and subjects were tested. In spite of this great variety of system elements, the predominant system closed-loop characteristic always fell within a very restricted range. For the type of plants being considered here, that is, plants of the form K/s and K/s^2, which have no static stability of their own, the predominant system characteristic was an oscillatory mode of motion. The frequency of this oscillation always fell within the range of from 4 to 2 rad/sec, and the oscillation was always stable. From these results, it is concluded that the pilot uses the predominant system characteristic as the criterion to set his static gain.

The predominant system characteristic is not always an oscillatory mode of motion. With the aircraft pitch response type of plant, the predominant system characteristic was a first-order factor. In the case mentioned before, the time constant of this first-order response was 1.7 seconds. There is also an oscillatory mode of motion included on the system response in this case, and it is also shown that this oscillatory mode of response is stable.

From these examples, it is concluded that the system response which the pilot will concern himself with can be of two different types. The first is the case when the predominant system characteristic is an oscillatory mode of motion; the pilot will be concerned with the frequency and damping of this mode of motion. In the second case, when the predominant system characteristic is a first-order mode of motion, the pilot will be concerned with the time constant of this mode and the stability of the other oscillatory modes of the response.

If the control task is a multiloop problem, where there is a second variable (such as altitude or horizontal displacement) which is controlled by the manipulation of an inner-loop variable (such as pitch angle), it is necessary to specify an outer-loop pilot model in a series with the inner-loop pilot. In reference 5, it is shown that in the multiloop control task of lunar module hover, where the outer-loop variable is horizontal translation and the inner-loop variable is pitch angle, the series arrangement of the pilot model does give a true representation of human pilot response. In this task the outer-loop dynamics were:

$$\frac{x}{\theta} = \frac{5.36}{s^2} \tag{4}$$

and the outer-loop pilot model was:

$$\frac{\theta_c}{x_e} = \frac{0.009(1 + 9.2s)}{(1 + 0.1s)^2} \tag{5}$$

It was also shown in reference 3 that the lag function included in the outer-loop pilot model is not really required to properly represent the pilot. The outer-loop variable, translation, is controlled at such a low frequency that a lag of the magnitude of 0.1 second does not influence the closed-loop response. Reference 6 has shown similar results for the case of hover with a helicopter. Again, a series arrangement was used to represent the pilot, and the outer-loop model was:

$$\frac{\theta_c}{x_e} = 0.027(1 + 0.4s)e^{-0.08s} \tag{6}$$

where a small time delay is used instead of a lag function. Reference 7 also used the series arrangement to represent the pilot in the case of altitude control for an airplane. In this case the outer-loop pilot was

$$Y_h = \frac{\theta_c}{h_e} = \frac{0.4}{1 + 0.1s} \tag{7}$$

In this case, no lead was required and again a very short time constant lag function was included.

In all three of these references (5, 6, and 7), the outer-loop variable is controlled at a lower frequency than is the inner-loop variable. Also, when the outer loop is closed around the inner loop, the resulting modes of motion couple with each other, and further restrictions on the outer-loop static gain can result from requirements to maintain stability of the short-period inner-loop variable.

Based on the results of these single-loop and multiloop experiments, different levels of preferred pilot response are proposed. For single-loop or inner-loop control, the levels are:

1. The pilot would prefer to operate as a simple amplifier with a lag time constant of 0.2 second. It is hypothesized that if suitable system response characteristics can be achieved with this type of pilot response, the system will be judged to be satisfactory.
2. If some compensation is required on the part of the pilot in order to stabilize the system, the compensation will take the form of lead, with a maximum time constant of 1 second. It is hypothesized that if suitable system characteristics can be achieved with this pilot response, the system will be judged acceptable but unsatisfactory.
3. If further compensation is required, it will be supplied by reducing the lag time constant to 0.05 second. It is hypothesized that if this compensation is required to stabilize the system, the system will be judged to be unacceptable.

The suggestion that a pilot would prefer to operate as a simple amplifier was made many years ago in reference 8. The present suggestion allows some lag in this response. This is not a trivial addition, as it greatly influences the stability of the system.

For the outer-loop response of the pilot, similar levels of preferred response are proposed. They are:

1. The pilot would prefer to operate as a simple amplifier. It does not appear to be necessary to specify any particular value for lag in this case, because the expected values of this lag will not influence system response.
2. If additional compensation is required, it will take the form of lead. It appears that outer-loop lead time constants can be quite large.

For all levels of pilot response, the static gain will be adjusted to provide the quickest response possible while maintaining stability.

Pilot Rating Data

A large amount of data on pilot ratings as a function of various aircraft parameters has been accumulated in the past. Perhaps the most fundamental relation is that established between pilot ratings and aircraft longitudinal short-period characteristics, ω_n and $2\zeta\omega_n$. The boundaries dividing pilot ratings of satisfactory from acceptable-unsatisfactory and acceptable from unacceptable as a function of ω_n^2 and $2\zeta\omega_n$ are shown on Figure 1 for a landing condition. These boundaries were taken from reference 9 for a landing condition with good (front side) drag characteristic. tics.

Pilot ratings have also been shown to be a function of T_{θ_2} where

$$\frac{1}{T_{\theta_2}} = \frac{M_{\delta_e} L_\alpha - L_{\delta_e} M_\alpha}{M_{\delta_e}} \approx L_\alpha \qquad (8)$$

The relationship has been defined in many ways. In reference 10, it is defined by the statement that pilot opinion will be the best when

$$T_{\theta_2}\ \omega_n = 4 \qquad (9)$$

Also, pilot ratings have been shown to be a function of control actuator dynamics. In reference 11, it is shown that if the elevator actuator is represented by a second-order transfer function, and the frequency of actuator characteristic response is less than 20 rad/sec, pilot ratings will be seriously degraded.

It is proposed, therefore, to show that pilot-airplane system characteristics obtained using the various levels of preferred pilot response defined in the previous section will correlate with the pilot ratings as a function of $\omega_n - 2\zeta\omega_n$, T_{θ_2}, and actuator dynamics response.

Aircraft and System Representation

The airplane is represented by the two-degree-of-freedom longitudinal equations of motion

$$\dot{\alpha} - \dot{\theta} = L_{\alpha}\alpha \tag{10}$$

$$\ddot{\theta} = Mq\,\dot{\theta} + M_{\alpha}\alpha + M_{\delta_e}\delta_e \tag{11}$$

and the relationship for altitude

$$\dot{h} = V(\theta - \alpha)$$

Two different values of L_α were investigated, viz.: $L_\alpha = 0.585$ and $L_\alpha = 1.3$. M_{δ_e} was arbitrarily assumed to be equal to 1.0. The values of Mq and M_α were adjusted to provide a wide range of airplane short-period characteristics ω_n^2 and $2\zeta\omega$. It is not necessary to state a given value for velocity, V. In this analysis, it is assumed that the pilot will always adjust his static gain so as to achieve a certain period for the system response. It can be seen that the overall outer-loop static gain for the pilot-airplane combination will be K_hV; and therefore it is always the product which will be given, and it will be assumed to be independent of the value of V.

For altitude control, the complete block diagram of the pilot-airplane combination is shown in Figure 2. As can be seen, the outer-loop pilot model is assumed to be just a static gain. For attitude control, the outer loop was discarded and just the inner loop was considered. The inner-loop pilot model lag time constant was always assumed to be 0.2 second and the lead time constant was either 0 or 1 second.

RESULTS

First, the altitude response of the pilot vehicle system was determined for a wide variety of airplane characteristics with no lead included in the inner-loop pilot model. The inner- and outer-loop pilot model static gains were adjusted to give the best possible stable system response. It was found that two different

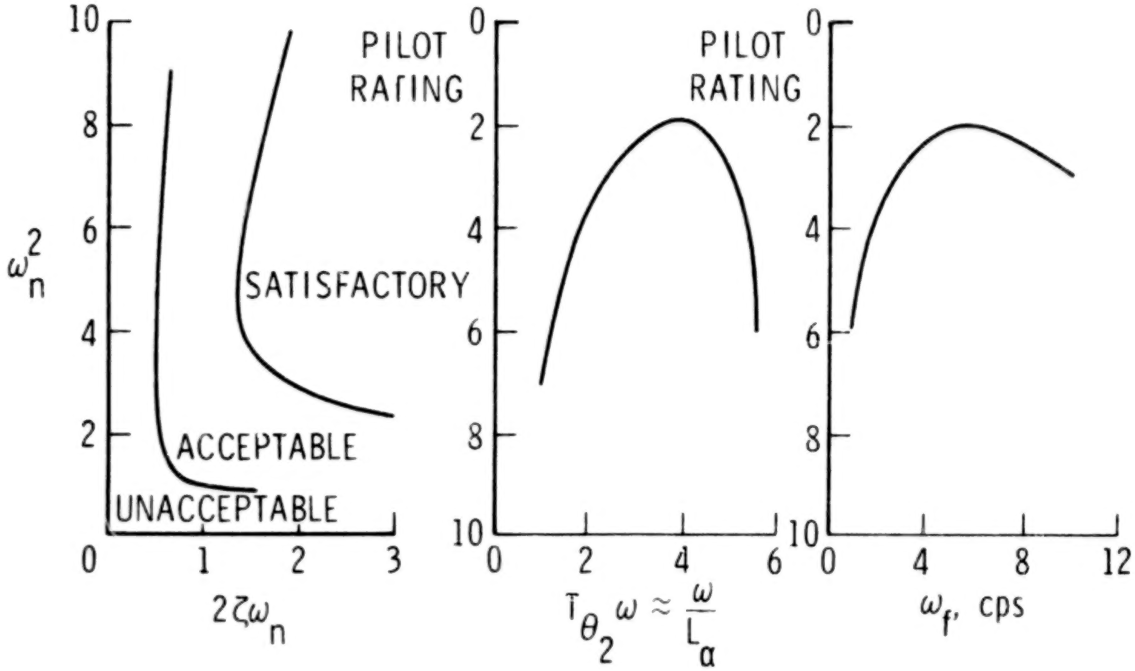


Figure 1.- Pilot ratings with short period ω_n - $2\zeta\omega_n$, L_α, and actuator frequency ω_f

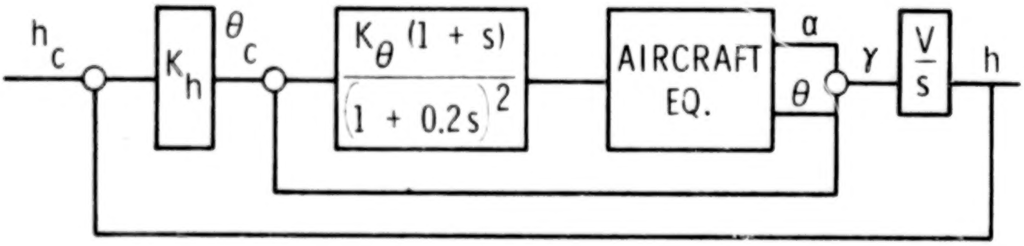

h_c
K_h
θ_c
$\frac{K_\theta(1+s)}{(1+0.2s)^2}$
AIRCRAFT EQ.
α
θ
γ
$\frac{V}{s}$
h

Figure 2.- System block diagram

definitions for best response were required. For the altitude control system, which was a sixth-order system, the response consisted of three oscillatory modes of motion. One of these modes was a very high-frequency, well-damped mode which corresponded to the control response, and it could not be considered as having a critical influence on the total system response. The other two oscillatory modes of motion corresponded to the angle-of-attack response and to the altitude response of the system. For airplanes with high values of ω_n^2, the angle-of-attack mode frequency was quite high, and therefore was not a critical factor. The critical factors were the altitude mode of motion frequency and damping and the damping of the angle-of-attack mode of motion. Therefore, the outer-loop and inner-loop pilot model's static gains were adjusted to as high a value as possible and still have a stable response, and the frequency of the altitude mode of motion was noted. For a value of $L_\alpha = 0.585$, typical values for the period of the altitude mode of motion are noted in Figure 3. The period is given because it is felt that the period can be more easily visualized, and thus will be more meaningful than will the frequency. The period of the angle-of-attack mode of motion is also given. It can be seen that for high values of ω_n^2, the altitude mode of motion period varied from 10 seconds when $2\zeta\omega$ was low to 4 seconds when $2\zeta\omega_n$ was high. From these results, it was decided that a period of 5 seconds would define a useful boundary to be used to divide the plot into satisfactory and acceptable areas. Accordingly, the synthesis method described in reference 12 was used to define this curve. It was specified that the altitude mode of motion period should be 5 seconds, and that the damping of the altitude mode of motion and the angle-of-attack mode of motion should be zero. The airplane ω_n^2 and $2\zeta\omega_n$ values that would satisfy these conditions were then determined; they are plotted on Figure 3. It can be seen that this curve agrees well with one branch of the pilot rating curve for a landing task.

For low values of ω_n^2, a second definition for a suitable response was required. For low values of ω_n^2, the angle-of-attack period became relatively long, and became the predominant and critical factor. Therefore, the pilot model static gains were adjusted to provide the shortest possible altitude and angle-of-attack periods, while maintaining a stable angle-of-attack response. The damping of the altitude mode of motion was always positive under these conditions, and therefore was not a critical factor. Typical values for the altitude and angle-of-attack periods are again shown in Figure 3. From these results and other considerations, it was decided that an angle-of-attack period of 2.5 seconds would provide a useful boundary to be used to separate the satisfactory from the acceptable region of the plot. The synthesis method was again used, where, this time, the altitude period was specified to be 5 seconds, the angle-of-attack period to be 2.5 seconds, and the angle-of-attack damping to be zero. The values of ω_n^2 and $2\zeta\omega_n$ that would satisfy those conditions are shown in Figure 3, and it can be seen that this curve agrees with the second branch of the pilot rating curve for the landing task.

Next, a 1-second lead time constant was added to the inner-loop pilot model, and system response was again calculated. With the addition of lead, the critical factors in the response were the period of the altitude mode of motion and the period and damping of the angle-of-attack mode of motion. When the criterion of 5 seconds for the altitude mode of motion period and 2.5-second period and zero damping of the angle-of-attack mode of motion was satisfied, the damping of the altitude mode of motion was always positive. The values of ω_n^2 and $2\zeta\omega_n$ which satisfy this criterion are shown on Figure 4, and it can be seen that these values agree fairly well with the pilot rating boundary between acceptable and unacceptable aircraft for a landing task. With the linear, constant coefficient pilot response used on these computations, there is no restriction in the lower and negative regions of ω_n^2.

Next, attitude control, with no pilot model lead, was examined. For this system, a fifth-order system, the system response consisted of two oscillatory modes of motion and a first-order mode of motion and was of no concern. For high values of ω_n^2, the frequency of the other oscillatory mode of motion was always high and therefore of no concern. The critical factors were the time constant of the first-order mode and the stability of the angle-of-attack mode. Therefore, the static gain of the pilot model was adjusted until the time constant was as low as possible with a stable angle-of-attack response. Typical values for these time constants are shown on Figure 5, and it can be seen that they vary from 6 to 2 seconds. From these results, it appeared that a value for the time constant of 2.6 seconds would provide a useful boundary. Again, the synthesis method was used to determine the values of aircraft ω_n^2 and $2\zeta\omega_n$ with which $T = 2.6$ and zero damping for the angle-of-attack mode of motion could be achieved. The resulting curve is shown on Figure 5 and can be seen to agree well with the pilot rating boundary between satisfactory and acceptable.

For lower values of ω_n^2, the period of the angle-of-attack mode of motion becomes a critical factor. The curve which satisfies the specification for an angle-of-attack period of 2.5 seconds and zero damping is shown on Figure 5, and can be seen to agree well with the other branch of the pilot rating curve. The time constant is always less than 2.6 seconds when these conditions are met.

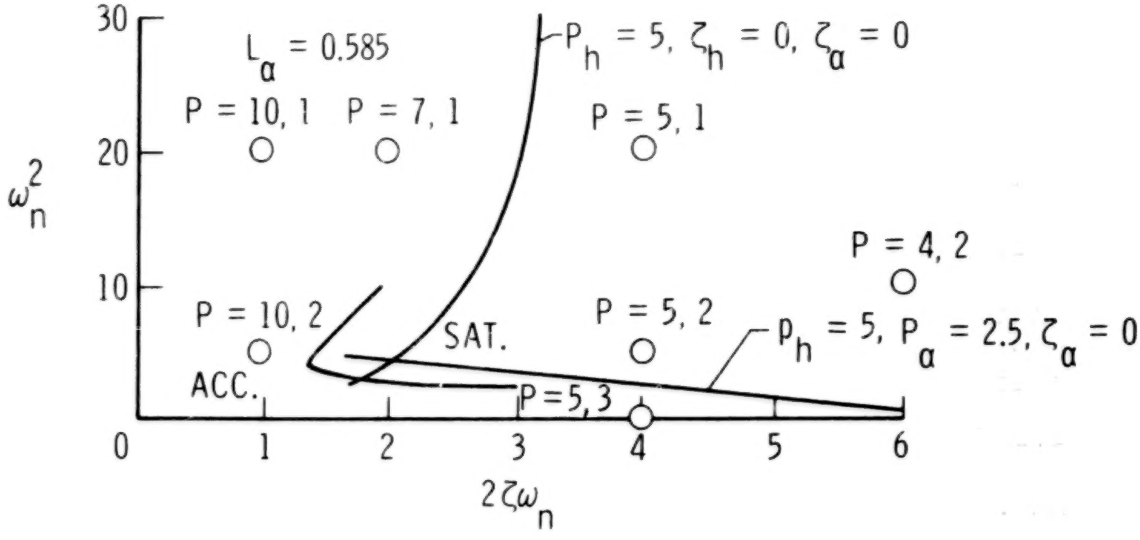


Figure 3.- Comparison of computed system characteristics using first level pilot models (no lead) for altitude control and pilot ratings

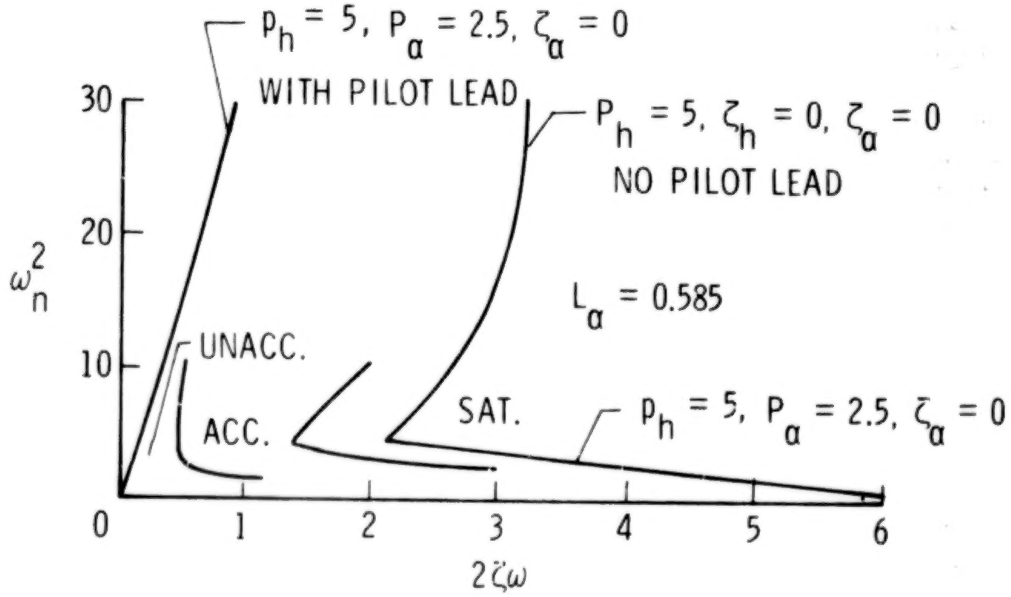


Figure 4.- Comparison of computed system characteristics using first and second level pilot models for altitude control and pilot ratings

From these results for altitude and attitude control, it is concluded that an aircraft which is rated satisfactory would be rated so on the basis of either altitude or attitude control tasks.

When lead is added to the pilot model for the attitude control task, the surprising result was that performance was degraded in the high ω_n^2 region (Figure 6). As was true in the case with altitude control, the addition of lead eliminated all restriction in the low and negative ω_n^2 region. However, instead of moving the boundary for satisfactory control in the high ω_n^2 region to the left, as was the result with altitude control, the boundary was moved to the right. This would seem to be a very confusing situation in which the addition of lead would help the pilot bring the aircraft under better control for altitude control, but degrades the system for attitude control. However, pilot comments indicate that an exceptional amount of confusion does exist for aircraft which fall in this region. In reference 13, page 128, pilot comments for an aircraft with the parameters $L_\alpha = 0.5$, $\omega_n = 7$ cps, $\zeta = 0.3$, which is indicated by the symbol on Figure 7, indicate that altitude control is fair, whereas attitude control is poor because of a tendency for the aircraft to bobble. It is felt that these comments are in good agreement with the computed results.

Another condition which can be used to test the validity of the pilot model is to see if they can confirm the change in pilot ratings that occur when L_α is changed. The effect of L_α on pilot ratings has been summarized by the statement that the pilot prefers that

$$T_{\theta_z} \omega_n \approx \frac{\omega_n}{L_\alpha} = 4$$

This formula indicates that for $L_\alpha = 0.585$, the preferred value for ω_n^2 is 5.5, which can be seen to correspond to the bulge in the curves for altitude and attitude control presented in Figures 3 and 5.

The boundaries for the same system characteristics, with no lead in the pilot model, were also calculated with L_α increased to 1.3. These results are shown, together with the results for $L_\alpha = 0.585$, in Figures 8 and 9. With $L_\alpha = 1.3$, the handling qualities data indicate that the preferred value for ω_n^2 increases to 27. The computed system characteristics indicate that the satisfactory area of the $\omega_n^2 - 2\zeta\omega$ plot has been greatly expanded in the high ω_n^2 region, and slightly restricted in the low ω_n^2 region for the higher value of L_α. Thus it can be seen that the computations agree with the pilot ratings.

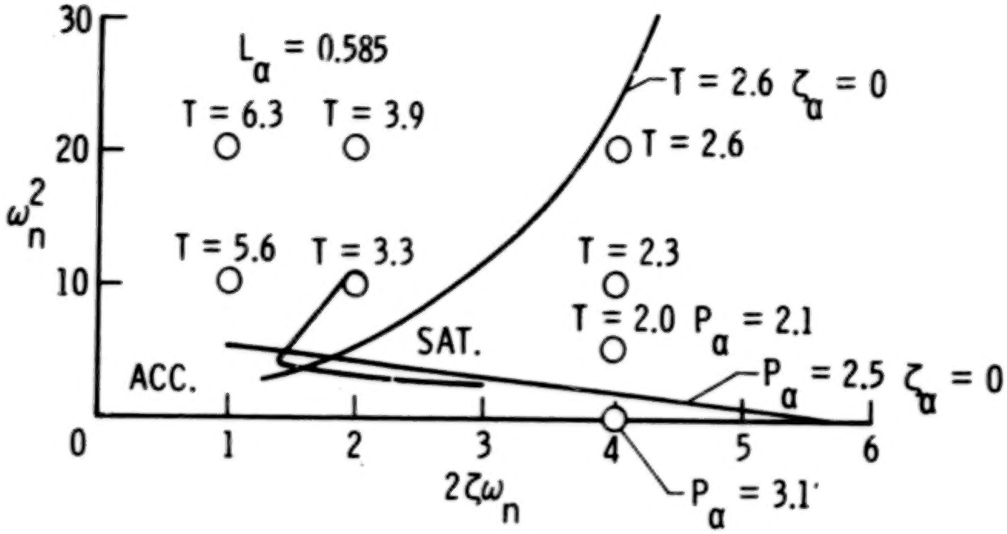


Figure 5.- Comparison of computed system characteristics using first level pilot models (no lead) for attitude control and pilot ratings

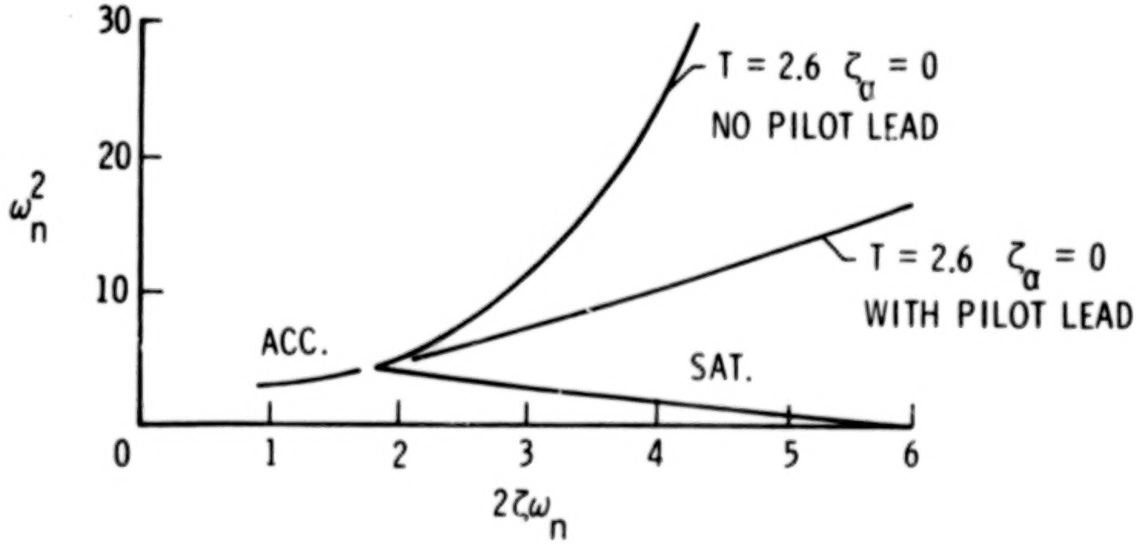


Figure 6.- Computed system characteristics using first and second level pilot models for attitude control

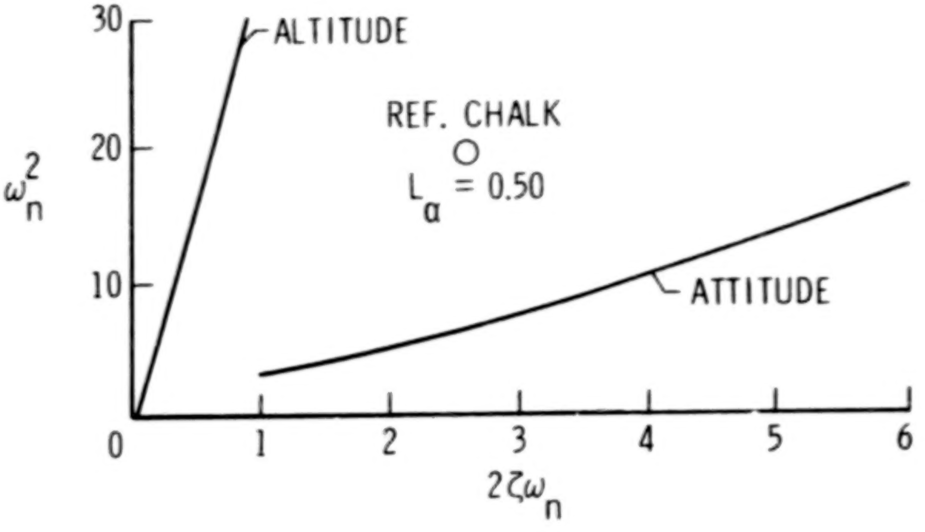


Figure 7.- Comparison of altitude and attitude control with second level pilot models

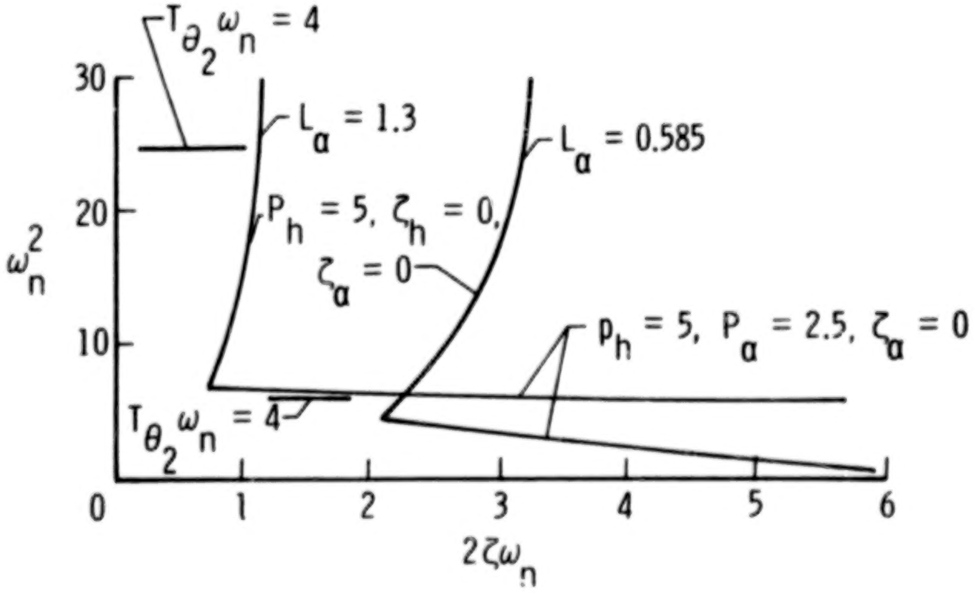


Figure 8.- Comparison of altitude control with first level pilot model for two values of L_α

A third influencing factor on pilot ratings that can be checked with the use of pilot models is the degrading effect of inserting control actuator dynamics in the control system. It has been shown that if the actuator dynamics have a natural frequency less than 20 rad/sec the pilot ratings will be drastically affected. Therefore, system characteristics were computed with actuator dynamics with a natural frequency of 10 rad/sec. This actuator would be expected to degrade the pilot rating from 2 to 4 numbers on the Cooper scale. The control situation examined was attitude control with no lead and with $L_\alpha = 0.585$. The results are plotted in Figure 10. As can be seen, the boundary for suitable system response is shifted to be more restrictive or, conversely, the response with a given aircraft is degraded by the addition of the actuator dynamics. For the good high-speed aircraft configuration tested in reference 11, which is plotted on the figure, the pilot rating is computed to change from satisfactory to acceptable. This change is in agreement with the pilot ratings presented in reference 11, where the rating changed from 3 to 5 with the addition of the actuator dynamics.

CONCLUSIONS

System response characteristics obtained using pilot models have been shown to correlate with pilot ratings in the following items:

1. In the shape of the boundaries in the airplane short period $\omega_n^2 - 2\zeta\omega$ plots.
2. In the change in pilot ratings with change in L_α.
3. In the change in pilot ratings with the insertion of control actuator dynamics.

It is therefore concluded that a prediction of pilot ratings can be made by determining the level of response of the pilot required to achieve specified system characteristics.

This method can also be used to compare competing designs and to provide a rational engineering interpretation to pilot comments.

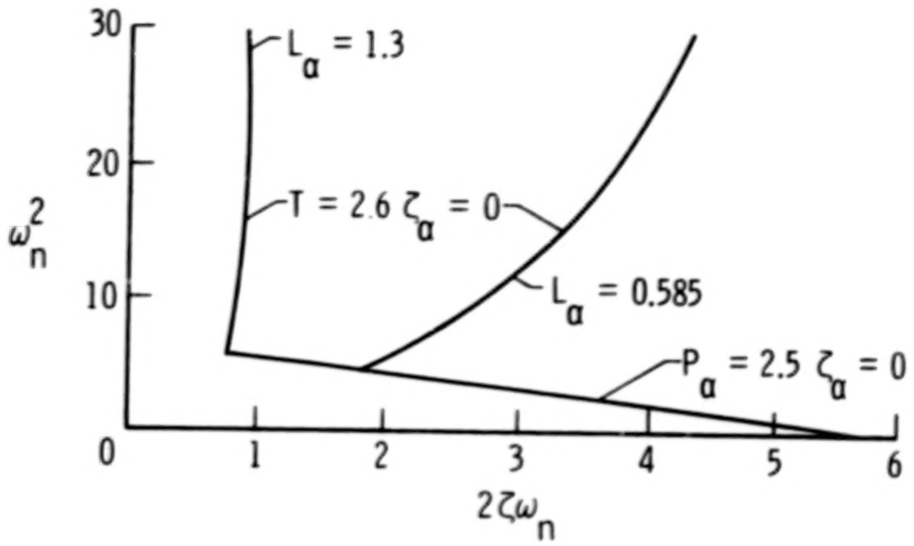


Figure 9.- Comparison of attitude control with first level pilot model for two values of L_α

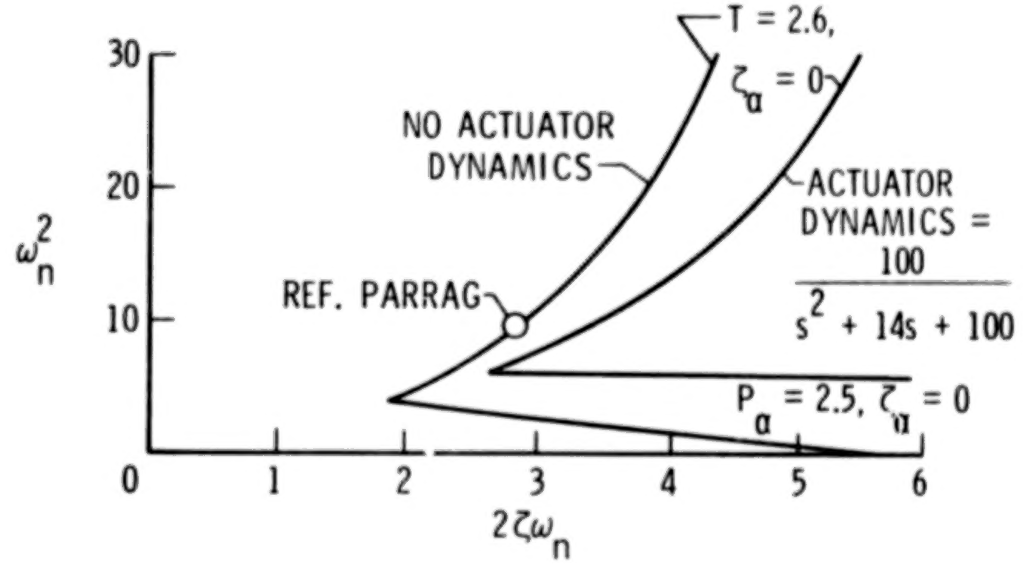


Figure 10.- Effect of actuator dynamics on calculated system response boundaries

REFERENCES

1. Adams, James J.; and Goode, Maxwell W.: "Application of Human Transfer Functions to System Analysis." NASA TN D-5478, 1969.

2. Adams, James J.; and Bergeron, Hugh P.: "A Synthesis of Human Response in Closed-Loop Tracking Tasks." NASA TN D-4842, 1968.

3. Adams, James J.; and Bergeron, Hugh P.: "Measured Variations in the Transfer Function of a Human Pilot in Single-Axis Tasks." NASA TN D-1952, 1963.

4. Adams, James J.; Kincaid, Joseph K.; and Bergeron, Hugh P.: "Determination of Critical Tracking Tasks for a Human Pilot." NASA TN D-3242, 1966.

5. Adams, James J.; Bergeron, Hugh P.; and Hurt, George J., Jr.: "Human Transfer Functions in Multi-Axis and Multi-Loop Control Systems." NASA TN D-3305, 1966.

6. Vinje, Edward W.; and Miller, David P.: "Interpretation of Pilot Opinion by Application of Multiloop Models to a VTOL Flight Simulator Task." NASA SP-144, 1967.

7. Stapleford, Robert L.; Craig, Samuel J.; and Tennant, Jean A.: "Measurement of Pilot Describing Functions in Single-Controller Multiloop Tasks." NASA CR-1238, 1969.

8. Birmingham, H. P.; and Taylor, F. V.: "A Design Philosophy For Man-Machine Control Systems." Proc. of the IRE, December 1954.

9. Chalk, Charles R.: "Flight Evaluation of Various Short Period Dynamics at Four Drag Configurations for the Landing Approach Task." FDL-TDR-64-60, 1964.

10. Hall, Warren G.: "In-Flight Investigation of Longitudinal Short Period Handling Characteristics of Wheel Controlled Airplanes." AFFDL-TR-68-91, 1968.

11. Parrag, Michael L.: "Pilot Evaluation in a Ground Simulator of the Effects of Elevator Control System Dynamics in Fighter Aircraft." AFFDL-TR-67-19, 1967.

12. Montgomery, Raymond C.; and Hatch, Howard G.: "Application of Differential Synthesis to Design of Multiaxis Stability Augmentation Systems." Journal of Aircraft, Vol. 6, No. 4, July-August 1969.

13. Chalk, Charles R.: "Fixed Base Simulator Investigation of the Effects of L_α and True Airspeed on Pilot Opinion of Longitudinal Flying Qualities." ASD-TDR-63-399, 1963.

A FREQUENCY DOMAIN APPROACH TO HANDLING QUALITIES DESIGN

By William A. Wolovich
Electronics Research Center

A. Introduction

The purpose of this paper is to introduce a new method for designing linear multivariable feedback control systems based on desired closed loop transfer matrix information. The technique which we employ to achieve the final design is based on a new theoretical result, known as the structure theorem. The structure theorem is a frequency domain relationship which simplifies the expression for the transfer matrix (matrix of transfer functions) of a linear time-invariant multivariable system. The effect of linear state variable feedback on the closed loop transfer matrix of the system is also clarified.

This paper is divided into seven sections. In section B, we discuss the vehicle, a Sikorsky SH-3D helicopter, and present the linearized equations of motion of the helicopter in a hovering position. We also precisely define linear state variable feedback, the method which we will employ to achieve a satisfactory control system.

The desired handling qualities are then discussed in section C. These handling qualities are first expressed qualitatively and then quantitatively, in terms of a desired closed loop transfer matrix. A practical system constraint, with implications regarding the maximum allowable feedback gains, is then presented in section D.

In section E, we present the main design tool, the structure theorem as it applies to a large class of practical systems, including the helicopter under investigation. We then proceed to design a linear feedback control system in section F using the information provided in the earlier sections. The final design is then compared to the desired design and some concluding remarks are made in section G.

We should state here at the outset that the design procedure which we employ is based on a prior knowledge of time-invariant, linearized equations of motion of a vehicle about some nominal point (in this paper, a hovering helicopter). No attempt has, as yet, been made to simulate or flight test the particular design presented in this paper. The results obtained, however, do appear promising, and further studies which will thoroughly evaluate the proposed design are planned.

B. The Mathematical Model

The vehicle which we will consider is the Sikorsky SH-3D Sea King, whose primary mission is to detect submerged submarines through the use of a sonar ball which is lowered into the water while the helicopter is hovering at an altitude of approximately 40 feet. We will begin our design process by considering the linearized equations of motion of this vehicle for small deviations from the trim condition. These equations were first presented in [1], and are given here in reordered form for reasons which we will discuss later, in section F.

The equations of motion are given in terms of the state-space representation;

$$\dot{\underline{x}} = \underline{A}\underline{x} + \underline{B}\underline{\mu} \ , \ \underline{y} = \underline{C}\underline{x}, \tag{1}$$

where $\underline{x}$ is an n-vector, called the state, $\underline{\mu}$ is an m-vector, called the input, and $\underline{y}$ is a p-vector, called the output. $\underline{A}$, $\underline{B}$, and $\underline{C}$ are constant matrices of the appropriate dimensions. For the SH-3D helicopter in hover, n = 9, m = 4, p = 6, and the normalized state $\underline{x}$, output $\underline{y}$, input $\underline{\mu}$, $\underline{A}$, $\underline{B}$, and $\underline{C}$ are defined as follows:

x_1; y_1 - longitudinal velocity

x_2; y_2 - pitch

x_3 - pitch rate

x_4; y_3 - vertical velocity

x_5; y_4 - roll

x_6 - roll rate

x_7; y_5 - yaw

x_8 - yaw rate

x_9; y_6 - lateral velocity

μ_1 - longitudinal cyclic pitch

μ_2 - main rotor collective

μ_3 - lateral cyclic pitch

μ_4 - tail rotor collective

Furthermore,

$$\underline{A} = \begin{bmatrix} -.016 & -.05 & .0025 & 0 & 0 & -.0001 & 0 & 0 & -.0047 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1.97 & 0 & -.542 & 1 & 0 & .548 & 0 & 0 & .736 \\ 0 & 0 & .00018 & -.3242 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 2.61 & 0 & -1.94 & -.163 & 0 & -1.96 & 0 & .01 & -7.25 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ .016 & 0 & -.0083 & -.193 & 0 & -.0043 & 0 & -.303 & 5.59 \\ .0047 & 0 & -.0024 & -.0007 & .05 & -.0025 & 0 & .0009 & -.033 \end{bmatrix} \quad (E1)$$

$$\underline{B} = \begin{bmatrix} .05 & .005 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -6.15 & .69 & 0 & 0 \\ 0 & -.424 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2.13 & 21.81 & .3475 \\ 0 & 0 & 0 & 0 \\ 0 & 5.12 & .174 & -7.48 \\ 0 & .01 & .05 & .022 \end{bmatrix} \quad (E2)$$

$$\underline{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (E3)$$

Given this open loop system, our objective is to design a "linear feedback" control system based on certain desired handling qualities and constrained by a practical consideration. In the context of this paper, linear feedback will be defined as the control law

$$\underline{\mu} = \underline{F}\ \underline{x} + \underline{G}\ \underline{w}, \qquad (2)$$

where $\underline{w}$ is an m-dimensional external input, and $\underline{F}$ and $\underline{G}$ are (mxn) and (mxm) constant matrices, respectively. Pictorially, we can represent a system compensated by linear state variable feedback as follows:

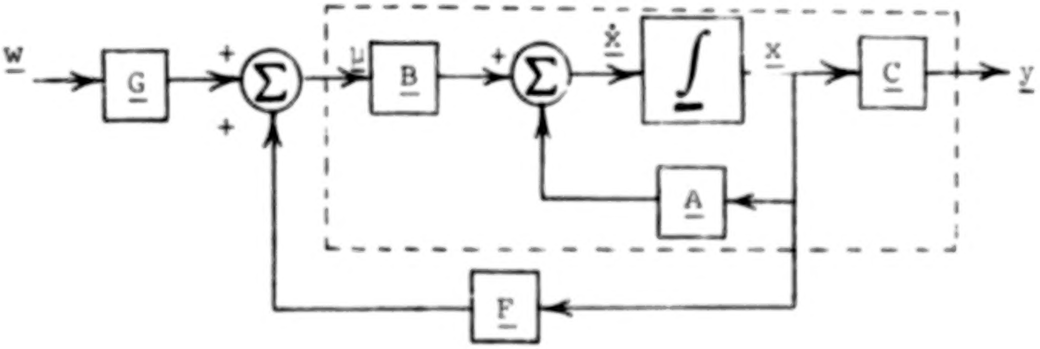


Figure 1.- Closed Loop System

The dotted portion of this block diagram represents the open loop system (1).

C. Desired Handling Qualities

We would like to design a control system in order to achieve certain desired handling qualities. More specifically, we would like to design a linear feedback control system which affords the pilot a relatively simple control task. This, in turn, implies a high noninteractive control system design; i.e. a design where each input affects only one or two selected outputs.

Therefore, we would like to achieve a final "decoupled" feedback design where longitudinal velocity and pitch (x_1, x_2, and x_3) are affected by longitudinal cyclic pitch (μ_1) alone. Simultaneously, it would be desirable if the main rotor collective input (μ_2) affected solely vertical velocity (x_4), the lateral cyclic pitch input (μ_3), affected only roll and lateral velocity (x_5, x_6, and x_9), and yaw (x_7 and x_8) was affected by tail rotor collective (μ_4) alone. In addition to this requirement, we would also like to select the poles associated with various input/output transfer functions as follows: Longitudinal velocity (y_1), lateral velocity (y_6), and yaw (y_5) should be affected by μ_1, μ_3, and μ_4 via pure integration. Furthermore, the transfer functions relating pitch (y_2) and roll (y_4) to longitudinal cyclic pitch (μ_1) and lateral cyclic pitch (μ_3) should represent second order systems with critical damping [2] ($\xi = .707$) and an undamped natural frequency [2] of 3 radians/second ($\omega_n = 3$). All other unspecified poles should lie in the half-plane Re $s<0$. These conditions represent the handling qualities which we hope to achieve using a linear feedback control system.

These somewhat qualitative handling qualities can be more precisely stated in the frequency domain in terms of a desired closed loop transfer matrix. In particular, if we solve for $\underline{y}(s)$, the Laplace transform of the output, in terms of $\underline{w}(s)$, the Laplace transform of the external input of our system, using (1) and (2), we obtain,

$$\underline{y}(s) = \underline{C}(s\underline{I} - \underline{A} - \underline{BF})^{-1}\,\underline{BG}\,\underline{w}(s), \tag{3}$$

were $\underline{C}(s\underline{I} - \underline{A} - \underline{BF})^{-1}\,\underline{BG}$ represents the closed loop transfer matrix of our system. For the SH-3D helicopter, the closed loop transfer matrix is a (6x4) matrix of transfer functions. Our design objective is to find a feedback control law, $\underline{\mu} = \underline{Fx} + \underline{Gw}$, if it exists, such that:

$$\underline{C}(s\underline{I} - \underline{A} - \underline{BF})^{-1}\,\underline{BG} = \begin{bmatrix} \frac{n_{11}(s)}{s(s^2+4.24s+9)} & 0 & 0 & 0 \\ \frac{n_{21}(s)}{s^2+4.24s+9} & 0 & 0 & 0 \\ 0 & \frac{n_{32}(s)}{s+1} & 0 & 0 \\ 0 & 0 & \frac{n_{43}(s)}{s^2+4.24+9} & 0 \\ 0 & 0 & 0 & \frac{n_{54}(s)}{s(s+10)} \\ 0 & 0 & \frac{n_{63}(s)}{s(s^2+4.24s+9)} & 0 \end{bmatrix} \tag{E4}$$

where the numerator polynomials $n_{11}(s)$, $n_{21}(s)$, $n_{32}(s)$, $n_{43}(s)$, $n_{54}(s)$, and $n_{63}(s)$ contain no pole cancelling zeros.

D. A Design Constraint

As stated in section B, there is a practical constraint associated with this design problem. In particular, if we consider the following block diagram of the SH-3D stabilization system, as given in [1],

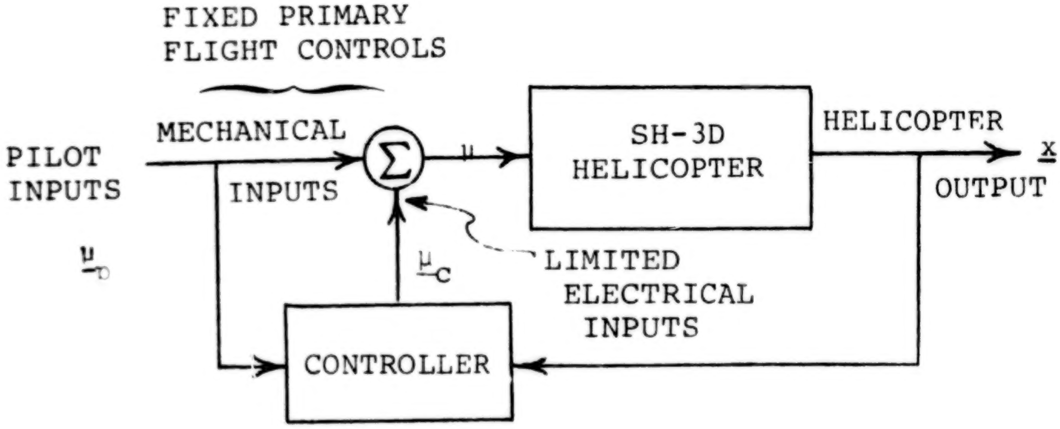

FIXED PRIMARY
FLIGHT CONTROLS
MECHANICAL
INPUTS
PILOT
INPUTS
$\underline{\mu}_p$
Σ
μ
SH-3D
HELICOPTER
HELICOPTER
OUTPUT
$\underline{x}$
$\underline{\mu}_c$
LIMITED
ELECTRICAL
INPUTS
CONTROLLER

Figure 2.- SH-3D stabilization system

we note that the control input, $\underline{\mu}$, consists of the pilot input, $\underline{\mu}_p$, and the controller output, $\underline{\mu}_c$. As stated in [1], the amplitude of the controller output must be limited to ± 10% of the total available range of the mechanical inputs. This limit enables the pilot to successfully recover from a hardover failure in the augmentation system. This, in turn, represents an amplitude limit on the entries of the feedback pair $\{\underline{F}, \underline{G}\}$, as we will now show.

The first step is to compare and equate figures 1 (section B) and 2. We note that figure 2 contains a direct forward path from the pilot input, $\underline{\mu}_p$, to the helicopter control actuators, $\underline{\mu}$. However, there is also an indirect path from $\underline{\mu}_p$ to $\underline{\mu}$ via the controller. Furthermore, there is a state variable feedback path from the helicopter output (assumed to be the entire state $\underline{x}$), to the controller. Mathematically, these factors imply that the controller output, $\underline{\mu}_c$, must satisfy the relationship,

$$\underline{\mu}_c = \underline{F}\underline{x} + \underline{G}^{*}\underline{\mu}_p \qquad (4)$$

Since the internal actuator control $\underline{\mu}$ is given by

$$\underline{\mu} = \underline{\mu}_p \pm \underline{\mu}_c, \tag{5}$$

we see that

$$\underline{\mu} = \underline{F}\underline{x} + (\underline{I} + \underline{G}^*)\underline{\mu}_p \tag{6}$$

Figures 1 and 2 are therefore "equivalent" if we set $\underline{\mu}_p = \underline{w}$ and $(\underline{I} + \underline{G}^*) = \underline{G}$. The entries of $\underline{F}$ and $\underline{G}$ (actually $\underline{G}^*$) must now be amplitude limited in order to satisfy the ± 10% controller output constraint.

A degree of engineering judgement is required at this point, since we do not have enough information to precisely state what gain limits should be imposed on the entires of $\underline{F}$ and $\underline{G}^*$. However, based on the final feedback design given in [1], we will require that all elements comprising the feedback matrix $\underline{F}$ be less than or equal to 2 in absolute value. Furthermore, based on the ± 10% constraint on $\underline{\mu}_c$, we will require that all elements comprising the input matrix $\underline{G}^*$ be less than or equal to 0.1 in absolute value. These numerical constraints on the entries of $\underline{F}$ and $\underline{G}^*$ will be imposed on the design procedure outlined in section F.

E. A Structure Theorem for Feedback Design

We present the main design tool in this section, a structure theorem which allows us to express the transfer matrix of a linear multivariable system in a concise and compact manner. The effect of linear state variable feedback on the closed loop transfer matrix is also clarified. The theorem which we will present here is useful only if one or more state variables are derivatives of other state variables. This condition occurs frequently in practice and does hold for the helicopter example considered. We should remark that a slightly different version of the structure theorem which we will present can be found in [3] and [4].

A number of definitions are required prior to stating the structure theorem. In particular, consider the system $\dot{\underline{x}} = \underline{A}\underline{x} + \underline{B}\underline{\mu}$. The pair $\{\underline{A}, \underline{B}\}$ is in multi-input companion form if, whenever $\dot{x}_i = x_j$, $j = i + 1$.[+] Suppose the pair $\{\underline{A}, \underline{B}\}$ is in

---

[+] In dealing with physical systems, this is usually the case. For the helicopter example considered, this condition clearly holds: i. e. $\dot{x}_2 = x_3$, $\dot{x}_5 = x_6$, $\dot{x}_7 = x_8$.

multi-input companion form. We partition $\underline{A}$ into the smallest possible number, q, of diagonal blocks, each a companion matrix of dimension σ_i, i = 1, 2, ...q. This partitioning of $\underline{A}$ for the helicopter example is demonstrated below. We see that $\bar{q} = 6$, $\sigma_1 = 1$, $\sigma_2 = 2$, $\sigma_3 = 1$, $\sigma_4 = \sigma_5 = 2$. and $\sigma_6 = 1$.

We now let $d_k = \sum_{i=1}^{k} \sigma_i$ for k = 1,2,...q. Clearly $d_1 = \sigma_1$, $d_2 = \sigma_1 + \sigma_2, \ldots d_q = n$.

$$\underline{A} = \begin{bmatrix} -.016 & -.05 & .0025 & 0 & 0 & -.0001 & 0 & 0 & -.0047 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1.97 & 0 & -.542 & 1 & 0 & .548 & 0 & 0 & .736 \\ 0 & 0 & .00018 & -.3242 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 2.61 & 0 & -1.94 & -.163 & 0 & -1.96 & 0 & .01 & -7.25 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ .016 & 0 & -.0083 & -.193 & 0 & -.0043 & 0 & -.303 & 5.59 \\ .0047 & 0 & -.0024 & -.0007 & .05 & -.0025 & 0 & .0009 & -.003 \end{bmatrix} \quad \text{(E5)}$$

We now define $\underline{A}_q$ as the (nxq) matrix consisting of the (q) ordered d_k - $\overline{\text{th}}$ rows of $\underline{A}$, and $\underline{B}_q$ as these same (q) ordered rows of $\underline{B}$. For the SH-3D helicopter,

$$\underline{A}_q = \begin{bmatrix} -.016 & .05 & .0025 & 0 & 0 & -.0001 & 0 & 0 & -.0047 \\ 1.97 & 0 & -.542 & 1 & 0 & .548 & 0 & 0 & .736 \\ 0 & 0 & .00018 & -.3242 & 0 & 0 & 0 & 0 & 0 \\ 2.61 & 0 & -1.94 & -.163 & 0 & -1.96 & 0 & .01 & -7.25 \\ .016 & 0 & -.0083 & -.193 & 0 & -.0043 & 0 & -.303 & 5.59 \\ .0047 & 0 & -.0024 & -.0007 & .05 & -.0025 & 0 & .0009 & -.033 \end{bmatrix} \quad \text{(E6)}$$

and

$$\underline{B}_q = \begin{bmatrix} .05 & .005 & 0 & 0 \\ -6.15 & .69 & 0 & 0 \\ 0 & -.424 & 0 & 0 \\ 0 & -2.13 & 21.81 & .3475 \\ 0 & 5.12 & .174 & -7.48 \\ 0 & .01 & .05 & .022 \end{bmatrix} \tag{E7}$$

Now define $\underline{S}(s)$ as the (nxq) matrix of monic single term polynomials in the Laplace operator s, i.e.

$$\underline{S}(s) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ s & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ s^{\sigma_1 -1} & 0 & & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & s^{\sigma_2 -1} & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & s^{\sigma_m -1} \end{bmatrix} \tag{7}$$

For the example considered,

$$\underline{S}(s) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & s & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & s & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & s & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{E8}$$

We also define

$$\underline{\delta}(s) = \text{diag}\,[s^{\sigma i}] - \underline{A}_q\underline{S}(s), \tag{8}$$

where

$$\text{diag}\left[s^{\sigma_i}\right] = \begin{bmatrix} s^{\sigma_1} & & & \\ & s^{\sigma_2} & & \underline{0} \\ \underline{0} & & \ddots & \\ & & & s^{\sigma_q} \end{bmatrix} \tag{9}$$

We can now state the structure theorem:

Theorem: *Consider the multi-input companion pair* $\{\underline{A}, \underline{B}\}$, *which defines the system* $\dot{\underline{x}} = \underline{A}\underline{x} + \underline{B}\mu$. *The open loop transfer matrix of the system*, $\underline{T}(s)$ *can be written as:*

$$\underline{T}(s) = \underline{C}(s\underline{I} - \underline{A})^{-1}\underline{B} = \underline{C}\underline{S}(s)\underline{\delta}^{-1}(s)\underline{B}_q, \tag{10}$$

where

$$|s\underline{I} - \underline{A}| = |\delta(s)| \tag{11}$$

Furthermore, *if we apply the linear state variable feedback control law*, $\underline{u} = \underline{Fx} + \underline{Gw}$, *then the closed loop transfer matrix*, $\underline{T}_{\underline{F},\underline{G}}(s)$, *can be written as*:

$$\underline{T}_{\underline{F},\underline{G}}(s) = \underline{C}(s\underline{I} - \underline{A} - \underline{BF})^{-1}\underline{BG} = \underline{CS}(s)\underline{\delta}_{\underline{F}}^{-1}(s)\underline{B}_q\underline{G} \qquad (12)$$

where

$$\underline{\delta}_{\underline{F}}(s) = \underline{\delta}(s) - \underline{B}_q\underline{FS}(s), \qquad (13)$$

and the characteristic polynomial of the closed loop system satisfies the relationship,

$$|s\underline{I} - \underline{A} - \underline{BF}| = |\underline{\delta}_{\underline{F}}(s)| \qquad (14)$$

Proof: For a detailed development and proof of these relationships and other related results, the reader should consult either [3] or [4].

F. The Design Procedure

We have now completely stated the design problem and discussed the technique which will be employed to seek an acceptable state variable feedback design. We will start the design process by recalling (12) and (13), which express the closed loop transfer matrix of the system in terms of the structure theorem; i.e. $\underline{T}_{\underline{F},\underline{G}}(s) = \underline{CS}(s)\underline{\delta}_{\underline{F}}^{-1}(s)\underline{B}_q\underline{G}$, where $\underline{\delta}_{\underline{F}}(s) = \underline{\delta}(s) - \underline{B}_q\ \underline{FS}$. For the helicopter example under consideration,

$$\underline{CS}(s) = \underline{I} \qquad (E9)$$

and $\underline{B}_q$ is given by (E7). Note that the elements which comprise the first and last rows of $\underline{B}_q$ are considerably smaller than those in the remaining (middle four) rows. Also recall that the elements of $\underline{F}$, the feedback matrix, are constrained in magnitude (section D). These two facts imply that linear state variable feedback will affect the middle four rows of $\underline{\delta}_F(s)$ significantly more than the first and last rows. Our design will therefore be based on properly altering the middle four rows of $\underline{\delta}_F(s)$ via $\underline{F}$. The effect on the other two rows of $\underline{\delta}_F(s)$ will be relatively insignificant, as we will show.

We stated in section B that the linearized equations of motion of the SH-3D helicopter, first presented in [1], are given here in reordered form. This was done so that there would be a numerically ordered correspondence between "coupled" input and output pairs; i.e. so that the desired closed loop transfer matrix (E4) would be diagonal in appearance. In particular, we would like the middle four rows of $\underline{T}_{\underline{F},\underline{G}}(s)$ to be strictly diagonal.

Now, note that for the example, $\underline{CS}(s) = \underline{I}$. $\underline{T}_{\underline{F},\underline{G}}(s)$ thus reduces to the product of three matrices; i.e.

$$\underline{T}_{\underline{F},\underline{G}}(s) = \underline{\delta}_{\underline{F}}^{-1}(s)\ \underline{B}_q\ \underline{G} \qquad \text{(E10)}$$

Furthermore, $\underline{F}$ and $\underline{G}$ (actually $\underline{G}^*$) can be altered, consistent with the constraint presented in section D. Our design will therefore be based on selecting $\underline{F}$ and $\underline{G}^*$ so that the product $\underline{\delta}_{\underline{F}}^{-1}(s)\ \underline{B}_q\underline{G} = \underline{T}_{\underline{F},\underline{G}}(s)$ is as diagonal as possible, consistent with the constraint that all of the elements, f_{ij}, of $\underline{F}$ satisfy the relationship,

$$|f_{ij}| \le 2, \qquad \text{(E11)}$$

and that $\underline{G}$ be of the form $\underline{I} + \underline{G}^*$, when all of the elements, g^*_{ij}, of $\underline{G}^*$ satisfy the relationship,

$$|g^*_{ij}| \le 0.1 \qquad \text{(E12)}$$

Also recall that we would like to arbitrarily assign all (9) poles of the closed loop system if possible (see section C).

We merely remark here that the only way the product $\underline{\delta}_{\underline{F}}^{-1}(s)\underline{Bg}\underline{G} = \underline{T}_{\underline{F},\underline{G}}(s)$ can be diagonal is if both $\underline{\delta}_{\underline{F}}^{-1}(s)$ and $\underline{B}_q\underline{G}$ are diagonal (actually we require that the middle four rows of these expressions be diagonal). The reader is referred to [4] for a thorough discussion of this point.

We now note that the middle four rows of $\underline{B}_q$, defined as $\underline{B}_{qm}$, represent a nonsingular (4x4) matrix; i.e.

$$\underline{B}_{qm} = \begin{bmatrix} -6.15 & .69 & 0 & 0 \\ 0 & -.424 & 0 & 0 \\ 0 & -2.13 & 21.81 & .3475 \\ 0 & 5.12 & .174 & -7.48 \end{bmatrix} \qquad \text{(E13)}$$

This, in turn, implies that $\underline{F}$ can be chosen such that,

$$\underline{\delta}_{\underline{F}}(s) = \begin{bmatrix} v & v & v & v & v & v \\ 0 & s^2+4.24s+9 & 0 & 0 & 0 & 0 \\ 0 & 0 & s+1 & 0 & 0 & 0 \\ 0 & 0 & 0 & s^2+4.24s+9 & 0 & 0 \\ 0 & 0 & 0 & 0 & s(s+10) & 0 \\ v & v & v & v & v & v \end{bmatrix} \tag{E14}$$

where the v's represent elements which have yet to be determined. In particular, if we let

$$\underline{F} = \underline{B}_{qm}^{-1} [\underline{A}_{qm} + \underline{F}_d] \tag{E15}$$

where $\underline{A}_{qm}$ is the (4x9) matrix consisting of the ordered middle four rows of $\underline{A}_q$, and

$$\underline{F}_d = \begin{bmatrix} 0 & -9 & -4.24 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -9 & -4.24 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -10 & 0 \end{bmatrix} \tag{E16}$$

then using (8), (9), and (13), $\underline{\delta}_{\underline{F}}(s)$ will equal the desired "diagonal" matrix (E14). For the example,

$$\underline{B}_{qm}^{-1} = \begin{bmatrix} -.1626 & -.2646 & 0 & 0 \\ 0 & -2.3585 & 0 & 0 \\ 0 & -.2045 & .0458 & .0021 \\ 0 & -1.62 & .001 & -.1336 \end{bmatrix} \tag{E17}$$

$$\underline{A}_{qm} = \begin{bmatrix} -1.97 & 0 & .542 & -1 & 0 & -.548 & 0 & 0 & -.736 \\ 0 & 0 & -.0002 & .3242 & 0 & 0 & 0 & 0 & 0 \\ -2.61 & 0 & 1.94 & .163 & 0 & 1.96 & 0 & -.01 & 7.25 \\ -.016 & 0 & .0083 & .193 & 0 & .0043 & 0 & .303 & -5.59 \end{bmatrix} \tag{E18}$$

and,

$$\underline{F} = \begin{bmatrix} .3203 & 1.463 & .6013 & .3414 & 0 & .0891 & 0 & 0 & .1204 \\ 0 & 0 & .0004 & 1.594 & 0 & 0 & 0 & 0 & 0 \\ -.1195 & 0 & .0888 & .1461 & -.4122 & -.1044 & 0 & -.0208 & .3202 \\ -.0005 & 0 & .0012 & 1.069 & -.009 & -.0029 & 0 & 1.296 & .7540 \end{bmatrix} \tag{E19}$$

We note that all elements of $\underline{F}$ satisfy the constraint (E11).[+] Now that $\underline{F}$ has been determined, we can compute $\underline{A} + \underline{BF}$ and $\underline{\delta}_{\underline{F}}(s)$ completely; i.e.

$$\underline{A}+\underline{BF} = \begin{bmatrix} 0 & -.0427 & .0331 & .025 & 0 & .004 & 0 & 0 & .0013 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -9 & -4.24 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -9 & -4.24 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -10 & 0 \\ -.0011 & 0 & .002 & .0491 & .0394 & -.0077 & 0 & .0284 & 0 \end{bmatrix} \tag{E20}$$

[+] If (E11) were not satisfied, an alternate $\underline{F}_d$ (E16) would have to be used, and we would not be able to achieve the desired closed loop pole configuration. [4]

and, by inspection,

$$\underline{\delta}_{\underline{F}}(s) =$$

$$\begin{bmatrix} s & -.0331s+.0427 & -.025 & -.004 & 0 & .0013 \\ 0 & s^2+4.24s+9 & 0 & 0 & 0 & 0 \\ 0 & 0 & s+1 & 0 & 0 & 0 \\ 0 & 0 & 0 & s^2+4.24+9 & 0 & 0 \\ 0 & 0 & 0 & 0 & s(s+10) & 0 \\ .0011 & -.002s & -.0491 & .0077s-.0394 & -.0284s & s \end{bmatrix} \tag{E21}$$

Hence,

$$\delta_{\underline{F}}^{-1}(s) \cong$$

$$\begin{bmatrix} \frac{1}{s} & \frac{.0331s-.0427}{s(s^2+4.24s+9)} & \frac{.025}{s(s+1)} & \frac{.004}{s(s^2+4.24s+9)} & 0 & \frac{.0013}{s^2} \\ 0 & \frac{1}{s^2+4.24s+9} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{s+1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{s^2+4.24s+9} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{s(s+10)} & 0 \\ \frac{-.0011}{s^2} & \frac{.022}{s^2+4.24s+9} & \frac{.0491}{s(s+1)} & \frac{-.0077s+.0394}{s(s^2+4.24s+9)} & \frac{.0284}{s(s+10)} & \frac{1}{s} \end{bmatrix} \tag{E22}$$

The final step in the design process involves $\underline{G}$; i.e. recall that for the example considered, $\underline{G} = (\underline{I} + \underline{G}^*)$, where each element of $\underline{G}^*$ must satisfy (E12). We would like to choose $\underline{G}^*$ such that the middle four rows of $\underline{B}_q\underline{G}$ represent a diagonal matrix. Observe that if $\underline{G}$ were equal to $\underline{B}_{qm}^{-1}$ (E17), then the middle four rows of $\underline{B}_q\underline{G}$ would equal the identity matrix. However, a number of the elements of $\underline{G}^*$ would clearly not satisfy (E12). If we normalize $\underline{B}_{qm}^{-1}$ by dividing each column by the corresponding diagonal element, and define this matrix as $\hat{\underline{B}}_{qm}^{-1}$, then

$$\hat{\underline{B}}_{qm}^{-1} = \begin{bmatrix} 1 & .112 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & .087 & 1 & -.016 \\ 0 & .687 & .022 & 1 \end{bmatrix} \qquad \text{(E23)}$$

If $\underline{G}$ were then set equal to $\hat{\underline{B}}_{qm}^{-1}$, two of the elements of $\underline{G}^*$ would equal .112 and .687 respectively, although this choice of $\underline{G}$ would also produce a diagonal $\underline{B}_{qm}\underline{G}$ matrix as desired. However, the element $g_{42}^* = .687$ would far exceed the constraint (E12), and it is therefore obvious that it is not possible to diagonalize $\underline{B}_{qm}\underline{G}$ and simultaneously satisfy the constraint (E12). The next best thing to do is to find a $\underline{G}$, if possible, which "almost" diagonalizes the product $\underline{B}_{qm}\underline{G}$; i.e. so that only one off-diagonal element of $\underline{B}_{qm}\underline{G}$ is nonzero. This is possible as we will now show.

In particular, suppose $\underline{G}$ were set equal to a modified version of $\hat{\underline{B}}_{qm}^{-1}$; i.e. Suppose

$$\underline{G} = \begin{bmatrix} 1 & .101 & 0 & 0 \\ 0 & .9 & 0 & 0 \\ 0 & .088 & 1 & -.015 \\ 0 & 0 & .022 & 1 \end{bmatrix} \qquad \text{(E24)}$$

Then

$$\underline{B}_{qm}\underline{G} = \begin{bmatrix} -6.15 & 0 & 0 & 0 \\ 0 & -.382 & 0 & 0 \\ 0 & 0 & 21.82 & 0 \\ 0 & 4.623 & 0 & -7.482 \end{bmatrix} \qquad \text{(E25)}$$

and

$$\underline{G}^* = \underline{G} - \underline{I} = \begin{bmatrix} 0 & .101 & 0 & 0 \\ 0 & -.1 & 0 & 0 \\ 0 & .088 & 0 & -.015 \\ 0 & 0 & .022 & 0 \end{bmatrix} \qquad \text{(E26)}$$

and, for all practical purposes, the elements of $\underline{G}^*$ would satisfy (E12); i.e., the only exception would be $g^*_{12} = .1\bar{0}1$ which can actually be set equal to .1 with negligible consequences. However, we will let $\underline{G}^*$ equal (E26) in our design. We can now compute $\underline{B}_q\underline{G}$ completely; i.e.

$$\underline{B}_q\underline{G} = \begin{bmatrix} .05 & .0096 & 0 & 0 \\ -6.15 & 0 & 0 & 0 \\ 0 & -.382 & 0 & 0 \\ 0 & 0 & 21.82 & 0 \\ 0 & 4.623 & 0 & -7.482 \\ 0 & .0134 & .0505 & .0213 \end{bmatrix} \tag{E27}$$

Our linear state variable feedback design is now complete; i.e. $\underline{F}$ is given by (E19), and the resulting $\underline{\delta}_F^{-1}(s)$ by (E22). $\underline{CS}(s) = \underline{I}$ and $\underline{B}_q\underline{G}$ is given by (E27). We can now employ (12) in order to obtain the closed loop transfer matrix for the hovering SH-3D helicopter, and then compare our final design to be the desired closed loop transfer matrix (E4);

i.e. $\underline{T}_{\underline{F},\underline{G}}(s) = \underline{CS}(s)\underline{\delta}_F^{-1}(s)\underline{B}_q\underline{G}$, and for the example,

$$\underline{T}_{\underline{F},\underline{G}}(s) = \left[\begin{array}{cccc|l} \frac{.05s^2 + .008s + .712}{s(s^2 + 4.24s + 9)} & \frac{.0096}{s + 1} & \frac{.0873}{s(s^2 + 4.24s + 9)} & 0 & \text{LONG VEL.} \\ \frac{-6.15}{s^2 + 4.24s + 9} & 0 & 0 & 0 & \text{PITCH} \\ 0 & \frac{-.382}{s + 1} & 0 & 0 & \text{VERT VEL} \\ 0 & 0 & \frac{21.82}{s^2 + 4.24s + 9} & 0 & \text{ROLL} \\ 0 & \frac{4.623}{s(s + 10)} & 0 & \frac{-7.482}{s(s + 10)} & \text{YAW} \\ \frac{-.135}{s^2 + 4.24s + 9} & \frac{.0134s^2 + .26s + .077}{s(s + 1)(s + 10)} & \frac{.051s^2 + .046s + 1.315}{s(s^2 + 4.24s + 9)} & \frac{.0213}{s + 1} & \text{LAT VEL} \\ \text{LONG. CYCLIC PITCH} & \text{MAIN ROTOR COLLECTIVE} & \text{LATERAL CYCLIC PITCH} & \text{TAIL ROTOR COLLECTIVE} & \end{array}\right] \tag{E28}$$

This closed loop transfer matrix, (E28), is proposed as the final design. If we compare this transfer matrix to the desired transfer matrix, (E4), several observations can be made. First we note that the middle four rows of (E28) closely resemble the middle four rows of (E4). The only exception is the (5.2) entry, i.e. the transfer function relating lateral velocity (y_5) to main rotor collective (u_2). This is not a completely desirable condition, although one can certainly tolerate this relatively low degree of interaction; i.e., when main rotor collective (u_2) is employed to change the vertical velocity (y_3), tail rotor collective (u_4) will have to be employed to counteract the resulting yaw (y_5).

We further note that with the exception of the (1,4) entry of (E28), all other elements in the first and last rows of the closed loop transfer matrix are nonzero. However, the gains associated with most of these transfer functions are relatively low; i.e. longitudinal velocity (y_1) is primarily controlled by longitudinal cyclic pitch (u_1), and lateral velocity (y_4) by lateral cyclic pitch (u_3), as desired. This desirable condition is not entirely due to our feedback design, however. Recall that the input gains associated with the first and last rows of $\underline{B}_q$ (E7) are low relative to the middle four rows of $\underline{B}_q$. Furthermore, lateral velocity and longitudinal velocity are "integrator coupled" to roll and pitch, respectively, in most aerodynamical vehicles. We simply took advantage of these factors in initially specifying a desired closed loop system (E4) based on realistic helicopter characteristics.

Although not all design objectives were met, it appears that this design is about as close to the desired design as possible, consistent with the constraints on $\underline{F}$ and $\underline{G}$. We should note here that if there were no constraint on $\underline{F}$ and $\underline{G}$, we could have achieved a strictly diagonal middle four rows of $\underline{T}_{F,G}(s)$, although the first and last rows would still closely resemble those of (E28).

G. Concluding Remarks

We have introduced the structure theorem and demonstrated its application to linear multivariable system design. In particular, we designed a noninteractive control system for a hovering helicopter using only pencil and paper methods.

Although the design outlined in this paper has not been simulated or flight tested, it appears to represent good handling qualities. The important point to note, however, is that the final design closely matches the desired closed loop transfer matrix.

We should remark here that the structure theorem outlined in this report is a modified version of a more general result which has been applied to areas other than handling qualities design. The reader should consult references [3] and [4] for additional information on this subject.

References

[1] Murphy, R. D. and Narendra, K. S. "Design of Helicopter Stabilization Systems Using Optimal Control Theory" J. Aircraft Vol. 6, No. 2, March-April, 1969, pp. 129-136.

[2] Truxal, J. G. Control System Synthesis, McGraw Hill, New York, 1955.

[3] Wolovich, W. A. and Falb, P. L. "On the Structure of Multivariable Systems" SIAM J. Control Vol. 7, No. 3, August 1969.

[4] Wolovich, W. A. A Frequency Domain Approach to the Design and Analysis of Linear Multivariable Systems Ph. D. Dissertation, Brown University, October, 1969.

Section 3

SYSTEM SIMULATION AND IDENTIFICATION

OPTIMAL REGULATOR OR CONVENTIONAL?
SETUP TECHNIQUES FOR A MODEL FOLLOWING SIMULATOR CONTROL SYSTEM

By Dwain A. Deets
Flight Research Center

SUMMARY

This paper compares the optimal regulator technique for determining simulator control system gains with the conventional servo analysis approach. Practical considerations associated with airborne motion simulation using a model-following system provided the basis for comparison. The simulation fidelity specifications selected were important in evaluating the relative advantages of the two methods. Frequency responses for a JetStar aircraft following a roll mode model were calculated digitally to illustrate the various cases. A technique for generating forward loop lead in the optimal regulator model-following problem was developed which increases the flexibility of that approach. In this study it appeared to be the only way in which the optimal regulator method could meet the fidelity specifications.

INTRODUCTION

Matching the motions of a pilot cockpit to those of an arbitrary aircraft is a fundamental problem of motion simulation. Two types of motion simulators are currently in use at NASA centers: ground-based and airborne. Many similarities exist between these two types of simulators. Both types are incapable of exactly duplicating all of the motions, and both employ complex multiaxis drive systems (Figure 1). Both types encourage the simulator engineer to ask-"With the same hardware, would the simulation be better if the control system were set up differently?"

This question led to the study reported in this paper. The simulator considered was the Flight Research Center's General Purpose Airborne Simulator (GPAS), a modified Lockheed JetStar (refs. 1 - 4). The GPAS utilizes a model-following system. The system gains for a roll-axis simulation were found using conventional methods, which normally would have been used in an actual flight simulation experiment. As an alternate approach, a different set of gains were found based on the optimal regulator technique, which is representative of modern control theory. Calculation of these gains by using each method was straightforward, but the determination of which simulation would have been better was difficult.

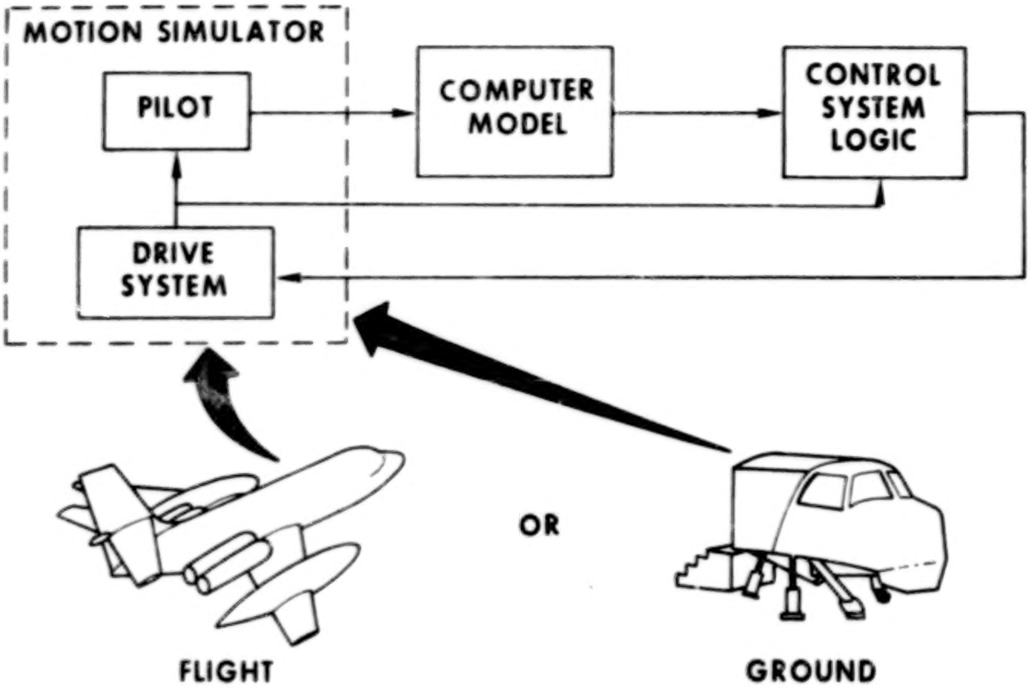


Figure 1.- Motion simulators

In order to make a sensible evaluation, ground rules were established which would be simple enough to follow for the study, but still close enough to the problems of motion simulation to be representative. Thus, the approach was to define a set of evaluation factors, including some representative specifications, on the simulation fidelity of the roll mode. Gains were then found by using conventional and optimal regulator techniques. The effort was to be limited to that which would normally have been expended in a typical simulation setup. After computing the gains and their respective characteristics, the various methods were to be evaluated against the original evaluation factors.

SYMBOLS

A	coefficient matrix for state equations
a	constant element in H_ℓ matrix
B	input matrix for state equations
e_p	error in roll rate, $(p_m - p)$
e_ϕ	error in bank angle, $(\phi_m - \phi)$
H	output matrix for state equations

J	optimal regulator performance index
j	imaginary axis of s plane
K	gain matrix
K_p	feedback gain, rolling angular velocity to aileron [sec]
K_ϕ	feedback gain, bank angle to aileron
p	rolling angular velocity [deg/sec]
Q	output weighting matrix
q_1, q_2, q_3	weighting values on e_p, e_ϕ, β, respectively
R	control weighting matrix
r	yawing angular velocity [deg/sec]
s	Laplace transform variable
t	time [sec]
$\underline{u}$	control vector
$\underline{x}$	state variable vector
$\underline{y}$	output error vector
β	angle of sideslip [deg]
δ_a	total aileron deflection, $\delta_{a_{left}} - \delta_{a_{right}}$, positive for left aileron -trailing edge down, [deg]
δ_{a_p}	pilot aileron command, positive when commanding positive δ_a [deg]
δ_r	rudder deflection, positive trailing edge left [deg]
σ	real axis of s plane
ϕ	bank angle [deg]

Subscripts:

m	model, analog-computer quantity
p	plant

Special notation:

(s)	function of Laplace variable
($\dot{\ }$)	a dot above a quantity represents first derivative with respect to time
$( )'$	transpose of a matrix or vector

EVALUATION FACTORS

Five evaluation factors were considered for application to this simulator design comparison. These were: pilot preference, time/facility requirements, simulation fidelity specifications, realizable system, and potential. Two of the factors, pilot preference and time/facility requirements, were discarded because they required information which could not be obtained within the scope of this study. The remaining three became the basis for evaluation. Brief comments on these three factors selected for the evaluation are given below.

Simulation Fidelity Specifications

One of the most significant evaluation factors for an analytical study is the system's ability to meet a given set of specifications. These requirements on the simulation fidelity have to be stated in terms of control system responses, but at the same time must reflect the pilot's threshold for distinguishing between various motion cues.

Realizable System

Another evaluation factor is the ability for a design procedure to automatically result in an achievable system. An example of a realizable system from a practical standpoint is one which requires relatively low valued feedback gains.

Potential

A final evaluation factor is the amount of potential a technique possesses. A method may exhibit potential in a number of ways. Easy expansion to more complex problems, ability to handle higher order effects, or the tendency for the method to lead to improved setup methods or improved simulation specifications are examples of "potential."

PROBLEM DESCRIPTION

A good technique for comparing design methods is to apply each method to a specific example. The ability of each method to cope with the example should be an indication of the merits or deficiencies to be expected with other similar problems. The roll mode problem described on the following page contains many specific details with regards to the model, simulator, gain configuration, and fidelity specifications.

Although the problem exhibited a few complications, such as having a multicontroller configuration, it was still relatively simple compared with the spectrum of motion simulator problems which could have been selected. The evaluation of the methods was not based solely on whether the specifications were met, but included the suitability of the method for handling motion simulation problems of greater complexity than this example.

Model-Following Configuration

A model-following system was used because it desensitizes the simulation to small changes which are likely to occur in the plant (JetStar) characteristics. The configuration studied is shown in Figure 2. The pilot's wheel input (δ_{a_p}) is fed into a prefilter model with a representative roll mode time constant of 0.35 sec. The equations for this model are written in state variable notation because of the ease of adapting to the optimal regulator problem format. Letting model roll rate (p_m) and bank angle (ϕ_m) be the model state variables, then

$$\dot{\underline{x}}_m = A_m\underline{x}_m + B_m u_m \tag{1}$$

with

$$\underline{x}'_m = [p_m \ \phi_m]; \ u_m = \delta_{a_p} \tag{2}$$

$$A_m = \begin{bmatrix} -2.86 & 0 \\ 1 & 0 \end{bmatrix}; \ B_m = \begin{bmatrix} 2.86 \\ 0 \end{bmatrix} \tag{3}$$

The plant was represented by the lateral-directional equations of motion of the JetStar at a cruise flight condition (Mach = 0.55, altitude = 20,000 ft, weight = 32,000 lb). In addition to roll rate (p) and bank angle (ϕ), sideslip angle (β) and yaw rate (r) were required to describe the lateral-directional dynamics. The ailerons (δ_a) and rudder (δ_r) were the controllers. The lateral-directional equations of motion of the JetStar are

$$\dot{\underline{x}}_p = A_p\underline{x}_p + B_p\underline{u}_p \tag{4}$$

with

$$\underline{x}'_p = [p \; \phi \; \beta \; r]; \; \underline{u}'_p = [\delta_a \; \delta_r] \tag{5}$$

and

$$A_p = \begin{bmatrix} -1.8 & 0 & -6.92 & .291 \\ 1.0 & 0 & 0 & 0 \\ .077 & .056 & -.14 & -1.0 \\ -.096 & 0 & 2.07 & -.28 \end{bmatrix}; \; B_p = \begin{bmatrix} 5.18 & 1.18 \\ 0 & 0 \\ 0 & -.003 \\ .153 & -1.79 \end{bmatrix} \tag{6}$$

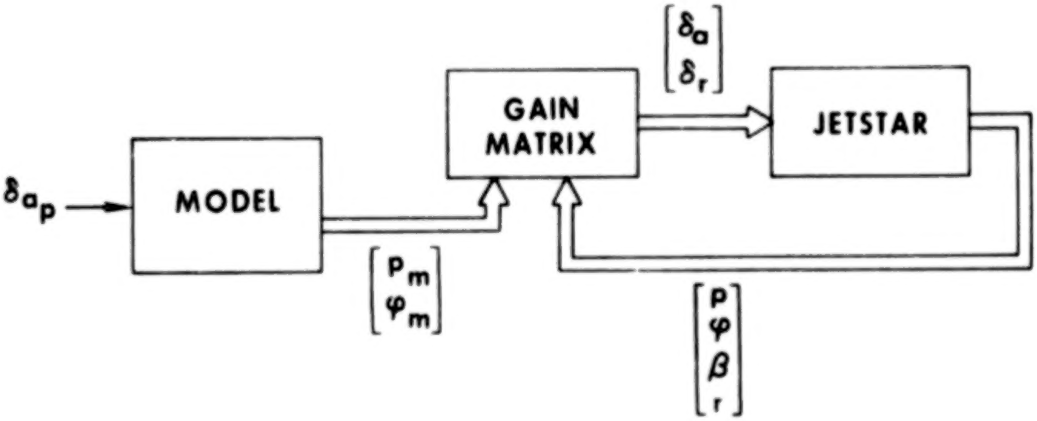

$\begin{bmatrix} \delta_a \\ \delta_r \end{bmatrix}$
GAIN MATRIX
JETSTAR
δ_{a_p}
MODEL
$\begin{bmatrix} p_m \\ \varphi_m \end{bmatrix}$
$\begin{bmatrix} p \\ \varphi \\ \beta \\ r \end{bmatrix}$

Figure 2.- Model-following configuration

The command to the control system was restricted to be a linear combination of model and plant state variables. Thus,

$$\underline{u}_p = K \begin{bmatrix} \underline{x}_m \\ \hdashline \underline{x}_p \end{bmatrix} \tag{7}$$

where K is the gain matrix noted in Figure 2.

The fact that the problems had to be formulated in state vector notation because of the requirements of the optimal regulator technique was viewed as being advantageous. Extension to much more complicated multiaxis problems would be straight forward with the optimal regulator, whereas conventional methods would become extremely cumbersome.

Roll Mode Simulation Specifications

The fidelity requirements for this study were stated as simply as possible. One requirement was that the lag in roll following should not be greater than 0.1 sec. This value was representative of specifications on previously conducted motion studies. The roll axis motion studies by Shirley (ref. 5) showed that a pure 0.1 sec time delay did not have a significant effect on the pilot's describing function in a tracking task.

Naturally, a 0.1 sec lag requirement is easy to achieve at low frequency inputs and difficult at higher frequencies. Thus, the specification required some form of frequency dependence. A 0.5 Hz pilot input was chosen as the typical pilot input for design purposes. Amplitude ratio was also recognized as an important fidelity measure. Very little was known about acceptable tolerances on motion amplitude ratios. A ±2 dB limit was chosen for the roll following.

These roll following specifications were treated as design goals for the study. It was expected to be extremely difficult or impossible to meet them with achievable system gains, and in fact they had not been met previously in simulations using the GPAS. In the directional mode, excursions in sideslip (β) were to be minimized with no specific tolerance.

The practical problems became apparent when restrictions were placed on the system gains. The most severe limitation on this problem was the need to use realizable roll rate and bank angle feedback gains to the JetStar's ailerons. Maximum allowable gains were selected based on previous flight experience with the JetStar's simulation systems.

Selection of gain maximums was actually a poor way to represent the real situation. Since the true gain maximums are due to factors such as structural or sensor modes rather than an absolute limit on the gain values, it would have been more beneficial to get at the source of the problems. However, this assumed that detailed knowledge of the aircraft structural modes, sensor dynamics, and all of the many interactions was available. Inclusion of this type of information was beyond the scope of this study. Therefore, limits on the gain values were used as the simplest method for making the system realizable. A summary of the specifications is found in Figure 3.

SIMULATION FIDELITY

TIME LAG AT 1/2 HERTZ INPUT 0.1 sec

AMPLITUDE RATIO TOLERANCE ±2 dB

REALIZABILITY CONSTRAINTS

MAXIMUM AILERON FEEDBACK GAINS

$K_p = 1 \qquad K_\varphi = 2.5$

Figure 3.- Motion simulator specifications
roll mode

DESCRIPTION OF DESIGN METHODS

Conventional Approach

Classical design methods were used to obtain the conventional gains. Considerable effort had been expended associated with developing the systems for GPAS leading to this conventional setup. Single loop Bode analyses which included actuator and sensor dynamics were used to estimate the gain requirements. Independent analysis was necessary for the aileron and rudder channels initially, then extended to the multiloop case using an iterative process. Ground and flight tests required various adjustments to the gains. This entire process resulted in a set of maximum usable gains. For the subject problem, the primary conventional gains were simply the maximum allowable values of the aileron feedback gains.

Optimal Regulator Approach

The state variable formulation of the system makes it easy to set up the optimal regulator problem. The model and plant are grouped together in a single matrix equation to give

$$\dot{\underline{x}} = \left[\begin{array}{c|c} A_m & 0 \\ \hline 0 & A_p \end{array}\right] \underline{x} + \left[\begin{array}{c} 0 \\ \hline B_p \end{array}\right] \underline{u} \qquad (8)$$

where

$$\underline{x}' = \left[\underline{x}'_m \,\middle|\, \underline{x}'_p\right]; \quad \underline{u} = \underline{u}_p \qquad (9)$$

Note that the model's input matrix (B_m) is not included in Eq. (8). This is because the optimal regulator solution assumes a response to arbitrary initial conditions on the state variables rather than an input from the pilot. This was potentially a problem in that accurate response resulting from a pilot's input is one of the most important qualities in a simulation. The fidelity specifications reflected this fact, since they were defined relative to a pilot's input. A question which had to be answered was whether a system, which had been optimized for response to initial conditions, would also perform satisfactorily for pilot inputs.

An output error vector ($\underline{y}$) consisting of pertinent error parameters was of the form

$$\underline{y} = H\underline{x} \tag{10}$$

where H is the output matrix.

For the roll following problem, the output vector was

$$\underline{y} = \begin{bmatrix} e_p \\ e_\phi \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \underline{x} \tag{11}$$

A quadratic cost functional (J) was defined to be

$$J = \frac{1}{2} \int_0^\infty (\underline{y}'Q\underline{y} + \underline{u}'R\underline{u})\, dt \tag{12}$$

where Q and R are positive semidefinite weighting matrices on the output error and control vectors, respectively. The optimal regulator problem minimizes the cost by finding a suitable control law. The control law obtained is of the form

$$\underline{u} = K\underline{x} \tag{13}$$

where K is a gain matrix.

The procedure for determining this gain matrix by finding the steady-state solution to a matrix Riccati equation is well documented in the literature (refs. 6, 7). Tyler (ref. 8) has studied specific applications to variable stability aircraft which were of direct applicability to this study. Tyler restricts Q and R to be diagonal matrices.

One guideline adopted for this study was that a diagonal weighting would be used unless there was a reason to do otherwise. For notational purposes, (q_1, q_2, q_3) stood for the Q matrix

$$Q = \begin{bmatrix} q_1 & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & q_3 \end{bmatrix} \qquad (14)$$

RESULTS AND DISCUSSION

Conventional Gains

Some of the characteristics of the conventionally determined gains are illustrated in Figure 4. The roll rate/pilot wheel $\left(\frac{p}{\delta_{a_p}}(s)\right)$ pole and zero locations are shown in the positive imaginary s plane. The JetStar's Dutch roll is well damped and canceled by a zero. Thus, directional quantities, (β and r), were not excited and did not contaminate the roll mode response. The model roll mode pole was a real root at -2.86. A highly damped complex pole pair and real zero combined to form a follower mode. This mode was the determining factor in the simulation fidelity.

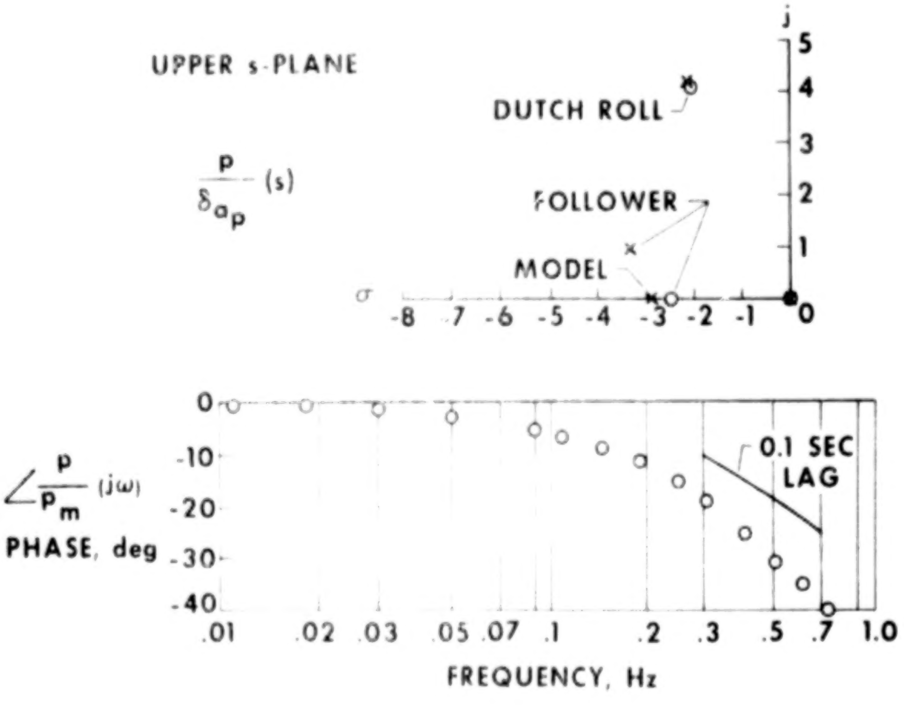


Figure 4.- Conventional gains

A phase angle frequency response of the $\frac{p}{p_m}$ (s) transfer function shows excessive lag, a strong indication that the design goal was to be difficult or impossible to meet. The amplitude ratio specification of ±2 dB was achieved out to 0.75 Hz.

Optimal Regulator Gains

The same problem was then solved using the optimal regulator procedure. In order to estimate the values of error weighting, five or six cases were run until the primary gains were close to the maximum limits. This weighting was (0.7, 7, 100) corresponding to (q_1, q_2, q_3). The time lag at 0.5 Hz was 0.2 sec, so another series of cases were run with equal e_p and e_ϕ weightings. The (1.5, 1.5, 100) case gave realizable gains. Some of the characteristics are shown in Figure 5. For reference, the conventional set is shown first. The gain values, follower mode roots, and the time lag at 0.5 Hz input are shown. Note that the follower mode for the (0.7, 7, 100) set has approximately the same characteristics as the conventional gains (as would be expected). The (1.5, 1.5, 100) set resulted in two real roots, with a zero canceling one of the poles. Unfortunately, the time lag was still longer than allowed by the specifications.

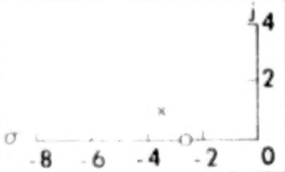
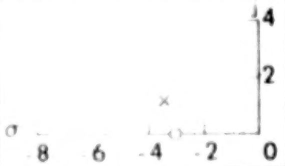
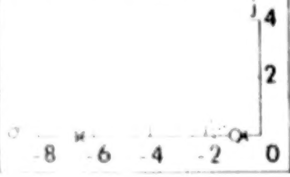
CASE	GAINS	FOLLOWER ROOTS	LAG
CONVENTIONAL	$\begin{array}{c|cccccc} & p_m & \varphi_m & p & \varphi & \beta & r \\ \hline \delta_a & 1 & 2.5 & -1 & -2.5 & 0 & 0 \\ \delta_r & 0 & 0 & 0 & -.11 & -10 & 2 \end{array}$		.17 sec
OPTIMAL (.7, 7, 100)	$\begin{array}{c|cccccc} & p_m & \varphi_m & p & \varphi & \beta & r \\ \hline \delta_a & .85 & 2.5 & -.96 & -2.5 & .8 & -.3 \\ \delta_r & .02 & -.02 & -.2 & -.15 & -8.4 & 2.9 \end{array}$		.20 sec
OPTIMAL (1.5, 1.5, 100)	$\begin{array}{c|cccccc} & p_m & \varphi_m & p & \varphi & \beta & r \\ \hline \delta_a & .93 & .91 & -1 & -.91 & .67 & -.23 \\ \delta_r & .03 & -.01 & -.24 & .16 & -8.4 & 2.9 \end{array}$		.17 sec

Figure 5.- Comparison between methods, conventional and optimal

Additionally, the amplitude fell below the -2 dB point at about 0.2 Hz and deteriorated with further increase in frequency. Although realizable gains were finally achieved, the procedure was by no means automatic. A large number of runs were necessary just to adjust the weightings to get the gains down to a reasonable value.

At this point in the study, the optimal regulator solution offered no advantage. In fact, it was not at all obvious how one could adjust the error weighting matrix (Q) such that the simulation fidelity would improve and still remain a realizable system. Conventional techniques did offer some straight forward methods using Bode analysis to adjust gains to give effective forward loop lead. Recognizing that forward loop lead was needed, the optimal regulator was examined for any method of introducing this lead in the problem.

Effect of Forward Loop Lead

Optimal regulator. - The optimal regulator performance index of Eq. (12) is repeated here for convenience.

$$J = \frac{1}{2} \int_0^\infty (\underline{y}'Q\underline{y} + \underline{u}'R\underline{u})\ dt \tag{12}$$

Now substituting in Eq. (10) for $\underline{y}$ gives

$$J = \frac{1}{2} \int_0^\infty [\underline{x}'(H'QH)\underline{x} + \underline{u}'R\underline{u}]\ dt \tag{15}$$

$H'QH$ is the state weighting matrix, being square with dimensions equal to the length of the state vector. Adjustment of the $H'QH$ matrix rather than the weighting matrix (Q) had the advantage of providing more flexibility and adjustments could be made that appealed to the designer's intuition.

Specifically, more forward loop lead was needed. This could be achieved by redefining the error for variables having derivatives which were state variables. The difference between the plants variable and a modified variable which leads the actual model was defined to be a model lead error. In the case of bank angle, the model lead error was

$$e_{\phi_\ell} = (as + 1)\phi_m - \phi \tag{16}$$

where "a" was an adjustable constant which acted like a free design parameter. Thus, a model lead error vector ($\underline{y}_\ell$) and a model lead output matrix (H_ℓ) for the roll mode simulation problem was of the form

$$\underline{y}_\ell = \begin{bmatrix} e_p \\ e_{\phi_\ell} \\ \beta \end{bmatrix} = H_\ell \underline{x} \tag{17}$$

where

$$H_\ell = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ a & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{18}$$

Five cases were run with increasing values of "a" from 0.1 to 5 for the (1.5, 1.5, 100) Q weighting matrix. The case for which "a" = 3 gave the best solution with respect to the simulation fidelity specifications.

The effect of the optimal regulator lead weighting can be seen in Figure 6. There has been a significant improvement in the $\frac{p}{p_m}(j\omega)$ phase lag at the lower frequencies. The amplitude ratio has increased as expected with the input lead, but is within the ±2 dB limit. The only difficulty is the failure to meet the time lag specifications at 0.5 Hz.

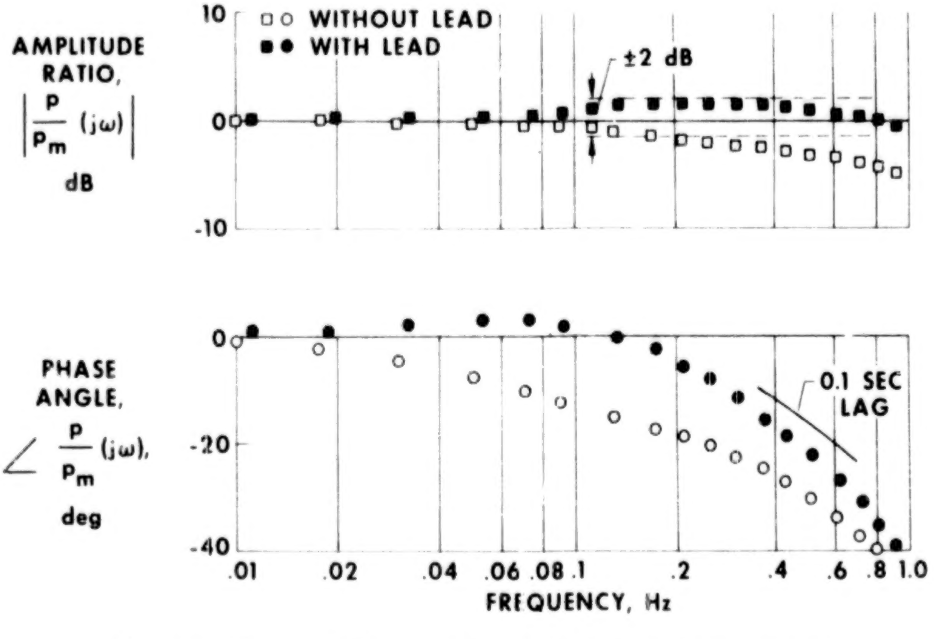


Figure 6.- Effect of optimal regulator lead weighting

The $H'QH$ weighting matrix for this case was

$$H'_{\ell}QH_{\ell} = \begin{bmatrix} 15 & 4.5 & -1.5 & -4.5 & 0 & 0 \\ 4.5 & 1.5 & 0 & -1.5 & 0 & 0 \\ -1.5 & 0 & 1.5 & 0 & 0 & 0 \\ -4.5 & -1.5 & 0 & 1.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (19)$$

and the control equation was

$$\begin{bmatrix} \delta_a \\ \delta_r \end{bmatrix} = \begin{bmatrix} 1.6 & .91 & -1.0 & -.91 & .67 & -.23 \\ .07 & -.01 & -.24 & -.16 & -8.4 & 2.9 \end{bmatrix} \begin{bmatrix} p_m \\ \phi_m \\ p \\ \phi \\ \beta \\ r \end{bmatrix} \qquad (20)$$

When compared with the basic (1.5, 1.5, 100) gains listed in Figure 5, the only gains to change were the p_m forward loop gains.

Adjustment of the $H'QH$ matrix to provide forward loop lead appears to be an important addition to the model following optimal regulator theory. It should have applications to control system problems other than for motion simulation considered here.

Conventional technique.- Comparison of the simulation fidelity provided by the optimal regulator lead weighting with the conventional technique could not be fairly made unless input lead adjustments were allowed on the conventional gains. Model lead was included using Bode diagram analysis. Increase in the p_m input gain to the aileron channel was required. The resulting characteristics were almost identical to the optimal regulator lead weighting case illustrated in Figure 6.

Final Comparison of Methods

With the completion of the simulator designs using each technique, the three evaluation factors specified earlier were used as a guide for comparison. Comments and an indication of the most suitable method for each evaluation factor are given below and summarized in Figure 7.

FACTORS	BEST METHOD
SIMULATION FIDELITY SPECIFICATIONS	EQUIVALENT
REALIZABLE SYSTEM	CONVENTIONAL
POTENTIAL	OPTIMAL REGULATOR

Figure 7.- Final comparison

Simulation fidelity specifications.- Both methods provided good simulation fidelity. Neither approach fully met all of the requirements, but the designer would have been in the position with either technique of choosing which tradeoffs should be made between the competing specifications. In the example studied, roll time lag was allowed to exceed the specifications slightly in order to meet the amplitude ratio and maximum feedback gain restrictions. The techniques were considered to be equivalent.

Realizable system.- The optimal regulator did not give achievable gains automatically. Information on hardware characteristics which actually limit the system could not be included in the optimal regulator easily, whereas it could be included as necessary in the conventional approach. The conventional technique was selected as the most suitable for providing a realizable system.

Potential.- The optimal regulator offered the most potential, considering its compact state variable notation and systematic handling of a multicontroller problem. Use of optimal control methods for setting up simulators would probably lead to much needed improvements in simulation fidelity specifications and simulator response documentation.

CONCLUDING REMARKS

The model-following optimal regulator technique, at this point in its development, does not appear to offer any strong advantages over conventional methods for determining the control law for simulator setup.

The most serious shortcoming of the optimal regulator method was the many iterations necessary to obtain a realizable system. Although state vector notation made the extension to more complex systems straightforward, the adjustments to the weighting matrix necessary for a realizable system would most likely become even more tedious and confusing.

An important output of this study was the development of a method for providing forward loop lead by using the optimal regulator solution without affecting the feedback gains. Without this new method, the optimal regulator would not have provided a simulation comparable to the conventional approach.

REFERENCES

1. Clark, Daniel C., and Kroll, John: "General Purpose Airborne Simulator-Conceptual Design Report." NASA CR-544, 1966.
2. Rediess, Herman A., and Deets, Dwain A.: "An Advanced Method for Airborne Simulation." J. Aircraft, vol. 1, no. 4, July-August 1964.
3. Kroll, John, Arendt, Rudy H., and Pritchard, Francis E.: "Development of a General Purpose Airborne Simulator." NASA CR-641, 1966.
4. Berry, Donald T., and Deets, Dwain A.: "Design, Development, and Utilization of a General Purpose Airborne Simulator." AGARD Rep. 529, May 1966.
5. Shirley, Richard S.: "Motion Cues in Man-vehicle Control." Massachusetts Institute of Technology, MIT-68-1, January 1968.
6. Athans, M., and Falb, P. L.: Optimal Control: An Introduction to the Theory and Its Applications. McGraw-Hill Book Co., Inc., New York, 1966.
7. Kalman, R. E., and Englar, T. S.: "A User's Manual for the Automatic Synthesis Program (Program C)," Chapter VII, NASA CR-475, June 1966.
8. Tyler, James S., Jr.: "Study of Models in Linear Optimal Control Theory and Application to Variable Stability Airplane Control System." CORNELL AERO LAB REPORT CAL-IM-1836-F-1, November 1967.

QUASI-OPTIMAL CONTROL OF A MOVING-BASE SIMULATOR

By Elwood C. Stewart
Ames Research Center

SUMMARY

This paper is concerned with the optimal control of a six-degree-of-freedom moving-base simulator, and utilizes the quasi-optimal control method due to Friedland. The problem, in broad terms, is to determine a control law for moving the simulator cab so that its motion will: (1) "best" approximate the desired aircraft response, and (2) not exceed the limited translational capability of the simulator. Due to lack of experimental data, "best" can only be interpreted subjectively at the present time. A variety of optimal responses have been obtained by the method which emphasizes different features of the response which are thought to be important in motion perception. Examples of such results are given, and the considerations important in initial subjective evaluations by experienced flight personnel are discussed. Final evaluation must be made on the motion simulation.

THE MOTION SIMULATOR PROBLEM

A photograph of the Ames six-degree-of-freedom motion simulator is shown in Figure 1. The purpose of the simulator is to provide motion cues to a pilot located in the cab by making it move so as to duplicate the motion of an actual aircraft. For this purpose the cab is provided with three rotational degrees of freedom with respect to the platform shown, and the platform is provided with three translational degrees of freedom. The lateral motion is well illustrated by the tracks on which the whole assembly moves. The motion of the cab is confined to a box roughly 8 ft longitudinally, 10 ft vertically, and 100 ft laterally. In addition to motion cues, the pilot is provided with visual cues by projecting on the screen in front of the cab the visual scene which would be seen from the actual aircraft. This visual information is obtained from another simulator.

The simulator control problem is illustrated in Figure 2. Here the pilot's input is sent to a model of the desired aircraft. Perfect visual cues are obtained from the aircraft output. The motion cues, however, are imperfect because they must be obtained from the simulator output rather than the aircraft. When the aircraft motions are small, the simulator motions can duplicate the aircraft motions so that the simulation will be 1:1. However, when the aircraft motions become large, the simulator cannot duplicate these motions and the simulation cannot be 1:1. Obviously a

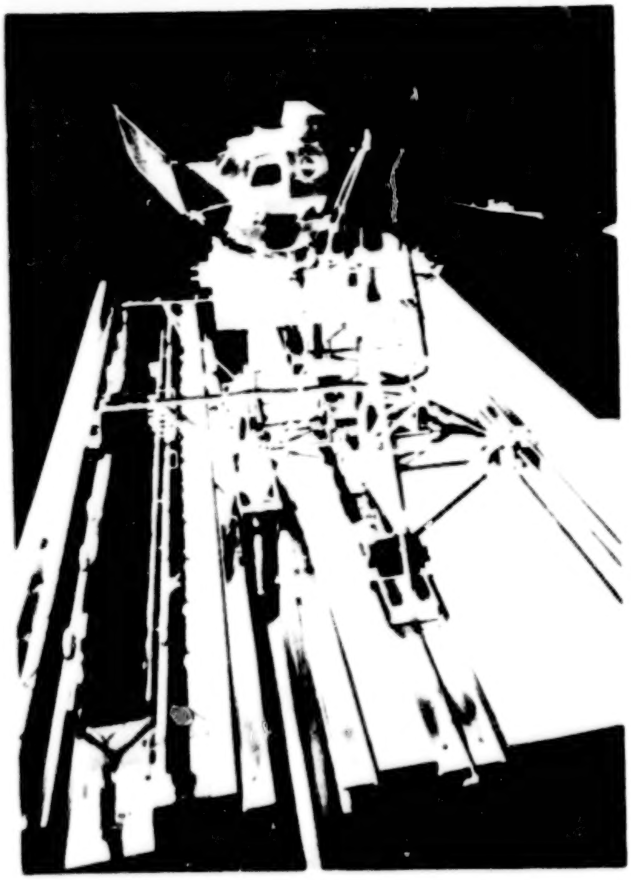


Figure 1.- Ames six-degree-of-freedom motion simulator

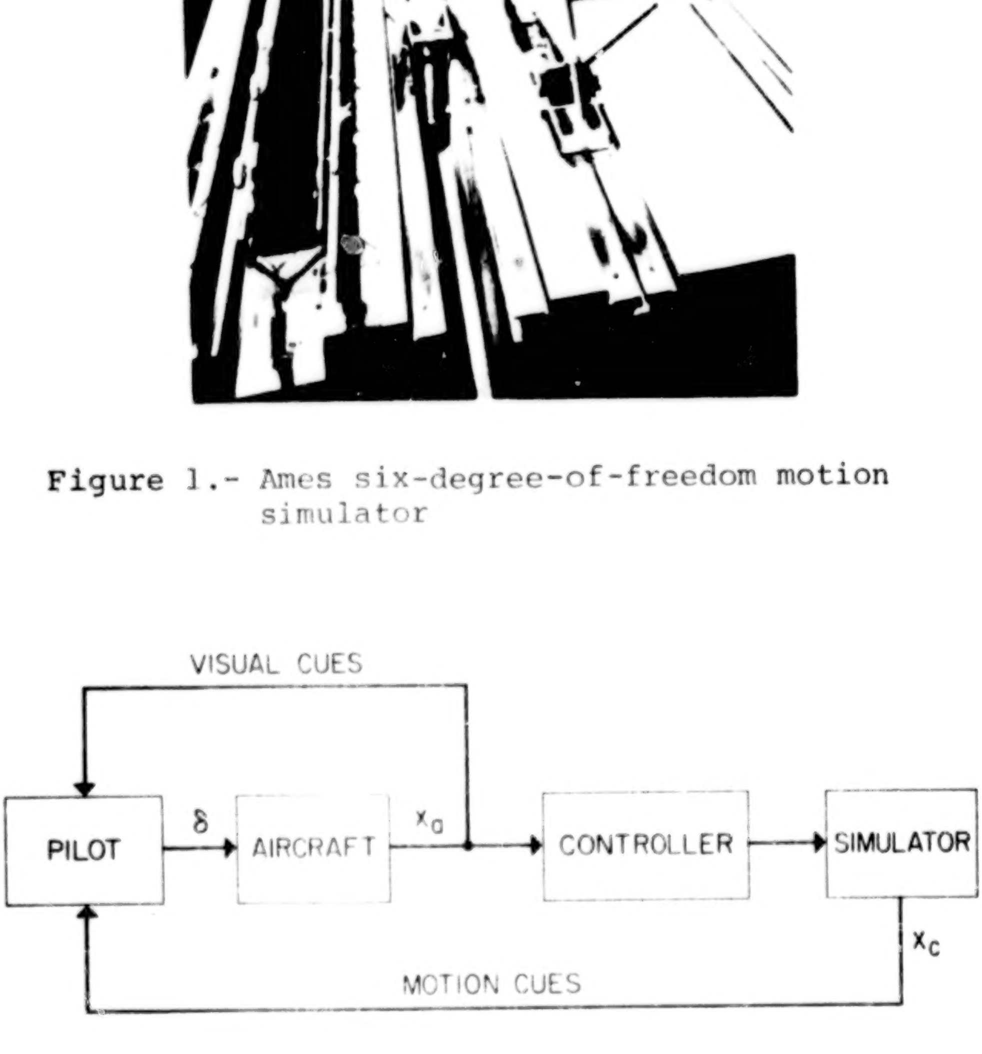

VISUAL CUES
PILOT
δ
AIRCRAFT
x_a
CONTROLLER
SIMULATOR
x_c
MOTION CUES

Figure 2.- Block diagram of simulator problem nonlinear controller is desired. The basic question in designing the control system is: What aspects of the aircraft motion must be followed accurately by the simulator and what aspects can be sacrificed without degrading the realism of the simulation?

Current methods of simulator control are based on using a high-pass filter as the controller in Figure 2. The approach seems to work satisfactorily for certain tasks. However, it is clear that the results obtainable are limited by the simple form of this controller; moreover, this type controller cannot provide the desired nonlinear behavior.

It is pertinent to examine an alternative approach utilizing optimal control theory. Such an approach would enable the desired nonlinear behavior to be achieved; further, better performance might be achieved; and finally, it might offer a more systematic approach to simulator controller design. The method being used is the quasi-optimal control method due to Friedland (refs. 1-3). A description of the method is beyond the intended scope here. It is sufficient to point out that the method is based on the Maximum Principle, but it avoids the usual two-point boundary-value problem by a unique perturbation idea. The study is being conducted by the combined efforts of three groups: first is Dr. Friedland and his group at Singer-General Precision (under contract to Ames Research Center); also involved in the study are the Theoretical Guidance and Control and the Flight and Systems Simulation Branches at Ames. As will be seen these three groups are required in the study because of the difficulty of defining the problem.

APPLICATION OF THE QUASI-OPTIMAL CONTROL METHOD

In attempting to apply the quasi-optimal control method to the control of the six-degree-of-freedom simulator, the scope of the investigation has been restricted for the present. First, only two degrees of freedom are being considered, vertical motion and pitch motion. Later results will be presented for a simple push-over maneuver such as would be executed to acquire a glide slope. Second, the equations of the aircraft and simulator are approximated by linear equations.

The heart of the problem lies in the definition of the performance function to be minimized. In Figure 3 the error has been defined as the difference between the aircraft and cab states. For realism in the simulation, some function of this error should be minimized by proper choice of the control system. Unfortunately it is not known what function of the error should be minimized because of lack of experimental data on motion perception. For this reason the best thing to do at the present time is to take a conservative view by minimizing an integrated squared error of

ERROR

$$e = x_a - x_c$$

PERFORMANCE FUNCTION

$$J = \int_0^T Q(\ddot{e}_x, \dddot{e}_x, \dot{e}_\theta, \ddot{e}_\theta)\,dt + k\int_0^T L(x_c)\,dt$$

SPECIAL FORMS

$Q_1(\ddot{e}_x, \dot{e}_\theta, \ddot{e}_\theta)$

$Q_2(\ddot{e}_x, \dddot{e}_x, \dot{e}_\theta, \ddot{e}_\theta)$

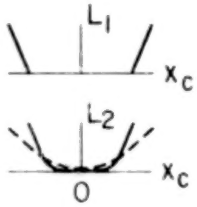


Figure 3.- Application of quasi-optimal control to six-degree-of-freedom simulator

those quantities which are known to be important in motion perception. It is well known from physiological experiments that the important quantities are linear acceleration, its derivative or jerk, angular rate, and angular acceleration as indicated in the performance function in Figure 3. The second part of the performance function is the penalty function term which is chosen so that L becomes large when the cab motion x_c approaches its limits of travel. This term has the effect of keeping the cab from reaching these limits. From this discussion it is clear that there is a certain freedom in the choice of Q and L as well as the constants in Q and L. Some special forms of interest are shown at the bottom of Figure 3. The constants in Q and L are much more difficult to choose, however. Any set of constants must satisfy two requirements. First, the displacement of the cab must not reach the limit of its travel. Second, the integral of the cab acceleration must be zero in order for the final cab velocity to be zero. Beyond that, the constants are more or less free to vary. In this study the constants have been purposely varied so as to generate a family of systems and responses, and an attempt is made by subjective evaluations to choose those which seem best. This is accomplished by a two-step selection process: (1) the eyeball test - from the time responses, experienced flight personnel are able to select subjectively those systems which appear to be best; (2) the flight or simulator test - here the promising systems are implemented on the six-degree-of-freedom simulator, and an experienced pilot flies the simulator in order to select the most satisfactory system. In this way it is hoped those characteristics of the response of most importance to motion perception can be determined.

The form of the optimal controller which is obtained by this method is shown in Figure 4. The controller is, in general, nonlinear, as desired. The controller requires information about the aircraft state, the simulator state, and an estimator of the pilot future control motions.

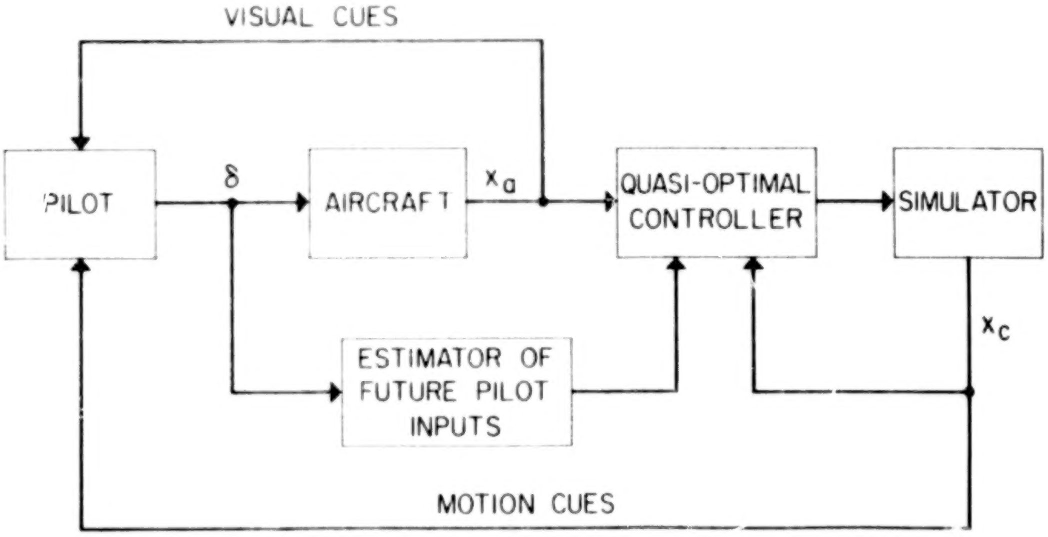


Figure 4.- Quasi-optimal control

Some typical results which are obtained by this method and the factors important in subjective opinions of them will now be discussed. Figure 5 shows a complete set of time histories which have been obtained for a system designed for one particular cost function where the particular maneuver is a simple push-over maneuver which would be required to go onto a glide slope for landing. It can be noted that the two requirements discussed above are satisfied; that is, the displacement remains less than 10 ft at all times and the final velocity is zero. It can also be noted that the errors in pitch and pitch rate are small. As far as pilot sensing is concerned, the pitch rate, its derivative, the linear acceleration, and jerk are important. Since the pitch-rate error is so small, the quality of the simulation will depend on the acceleration and jerk responses. However, it is not known in what way this quality depends on the errors in acceleration and jerk. Thus, as indicated previously, the cost function is purposely varied to generate a variety of acceleration of jerk responses. The responses for four different systems are shown in Figure 6. The displacement, velocity, pitch angle, and pitch rate are not shown, since they are similar to those shown in Figure 5. The two sets on the left are for Q_1, L_2 systems, that is, without jerk weighting but with a soft penalty function. The first set on the left is the same as in Figure 5, and it is generally believed that the acceleration response suffers from a poor phase relationship; that is, the zero crossing and peak times are in error.

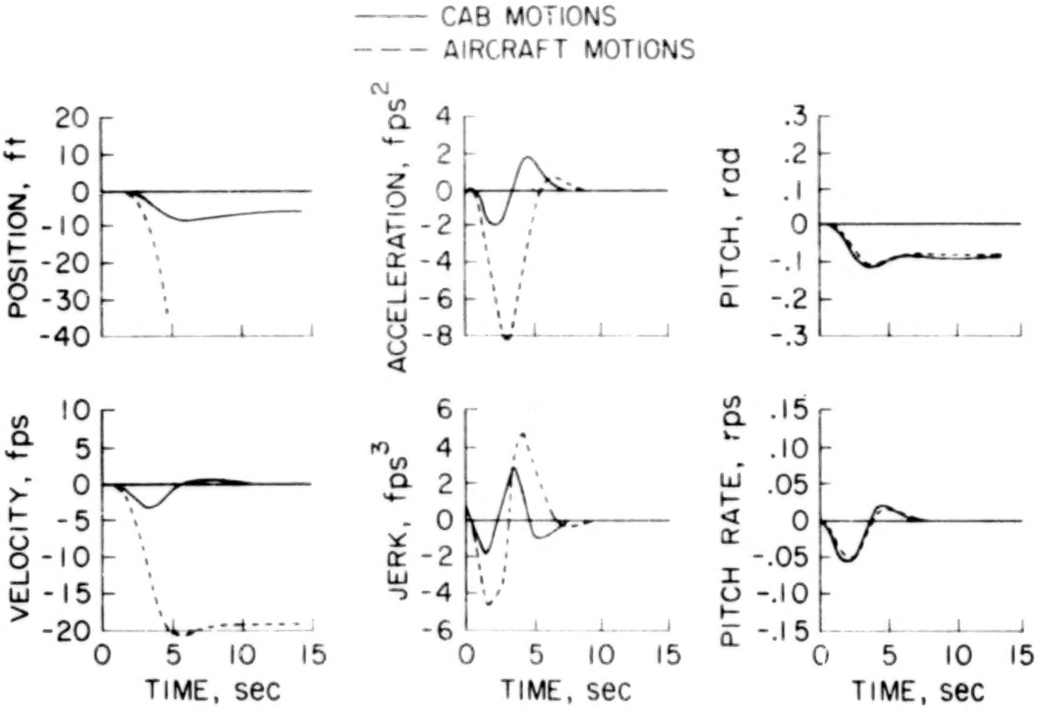


Figure 5.- Responses for example quasi-optimal control

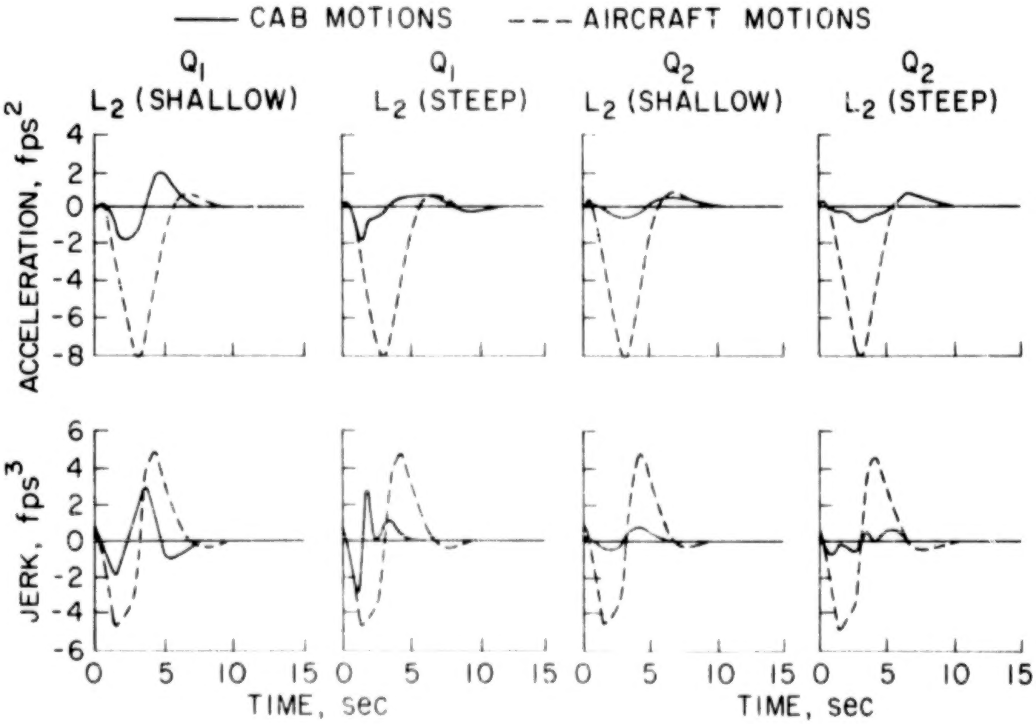


Figure 6.- Responses for quasi-optimal control with various weightings

In the second set, the sharpness of the penalty function has been increased, and, as a consequence, it is seen that the initial onset of acceleration is much more accurate. Some believe that this initial onset is of great importance in motion simulation. Note that the phase is still in error, however. The third and fourth sets are for Q_2, L_2 systems, that is, with jerk weighting and a soft penalty function. In the third set, the jerk weighting has greatly improved the phase relation of the acceleration response. However, the magnitude has been reduced and the onset of acceleration lacks fidelity. In the fourth set, the penalty function sharpness has been increased in order to provide a better onset of acceleration. But it can be noted that the jerk has developed some oscillation which might be undesirable.

The responses on this figure illustrate the variety of features which can be achieved by suitable choice of the performance function. The magnitude and phasing of the acceleration as well as the initial onset or jerk seem to be the three most important factors. However, which features prove to be the most important, and hence which control system is best, will be answered in future tests in which experienced pilots fly these systems on the motion simulator.

REFERENCES

1. Friedland, B.: "A Technique of Quasi-Optimum Control." ASME Trans., ser. D, J. Basic Engr. 88, 1966, pp. 437-443.

2. Friedland, B., Thau, F. E., Cohen, V. and Ellis, J.: "Study of Quasi-Optimum Feedback Control Techniques." NASA CR-527, 1966.

3. Friedland, B., Thau, F. E., Welt, S., Ling, C. K., and Shilder, M.: "Additional Studies of Quasi-Optimum Feedback Control Techniques." NASA CR-1099, 1968.

BLANK PAGE

SOME MODERN CONTROL TECHNIQUES FOR HUMAN OPERATOR MODELING AND IDENTIFICATION

By Richard S. Shirley
Electronics Research Center

SUMMARY

In this paper some recent results in quantitative modeling of the human operator are reviewed as well as a method for processing his input-output data. Modern control techniques are applied to the pilot operating in a dynamic compensatory tracking task. Trained pilots are known to behave in an optimal manner under fully stressed conditions. This has led to the development of an optimal control model of the human operator which takes his limitations into account. A rationale for determining the appropriate cost functions is determined. Experimental results indicate that the model can closely predict pilot performance in one- or two-axis tracking tasks. The model predicts pilot remnant (noise), error scores, and frequency response in both single-axis and multi-axis tracking tasks.

The model is evaluated by comparison with actual experimental data. The recently developed B. L. Ho procedure yields a linear system realization from the system impulse response or transfer function without assuming a priori the form of the system. A modification of the procedure is described in the paper,and actual experimental data are processed for a human operator in a single-axis compensatory tracking task, controlling dynamics which are (in one case) roughly equivalent to those of the lateral position of a hovering helicopter. The resulting frequency response of the system realization closely approximates the frequency response of the human operator (in the region of crossover) as measured by the method of Fourier coefficients.

INTRODUCTION

In this paper a model of the human operator and a method for processing his input-output data are reviewed. The model and data-processing method permit a greatly improved means of analyzing a man-vehicle system operating under instrument flight rules. Specifically, they offer the following:

1. The ability to predict the human operator's frequency response in a multi-axis compensatory tracking task;

2. The ability to predict the pilot's remnant and system error scores along with the frequency response; this is particularly important because the pilot's remnant power often exceeds 50 percent of the total control stick power;

3. A method of obtaining a linear system realization of desired order of the pilot from his input-output data without assuming model form.

The techniques described are dependent upon advances in mathematics, psychology, systems simulation, and digital computer technology. Let us briefly examine how these four fields have interacted with the history of manual control and led to the analysis techniques to be described.

Originally manual control was a combination of experiment and experience. Given a steam-driven car and a man, it was merely a matter of common sense to design levers which allowed the man to control the car. By trying various levers (experiment) improvements were found which were incorporated in future cars (experience). The system worked very nicely: given a reasonable set of control devices, the man could control the vehicle (barring mechanical failure). The coming of the airplane brought changes. For the first time engineers wanted to analyze the man-vehicle system before the experiment, for the cost of a poor design could be high; for instance, many pilots died flight-testing aircraft. Furthermore, new instruments were required beyond those monitoring the mechanical performance of the vehicle, for a pilot out of sight of land required navigational aids. It was at this point that psychologists first made a contribution to manual control, soon to emerge as "human factors." Psychologists aided engineers in the design of displays which could be easily seen and controls which could be easily used. The psychologists' concern with visual angles and scanning patterns resulted in a systematic approach to designing the man-vehicle interface.

During the 1940's and 50's, the concept of handling qualities emerged because of three developments. First, considerable flight experience had been accumulated, and rightly or wrongly, engineers and pilots had definite ideas of the kind of aircraft dynamics a pilot could control. Secondly, larger and faster aircraft were being built with dynamics which did not fit this category of "controllable aircraft." Finally, in order to cope with the larger, more complex systems, engineers began to adopt certain mathematical analysis procedures, such as frequency response analysis, root loci, Nyquist diagrams, and other classical techniques of control engineering. Their use permitted engineers to change the dynamics of the new aircraft to conform with their ideas of "controllable," giving the concept of handling qualities some value. The use of handling qualities has been refined until today numbers (the Cooper scale, for example) can be associated with various ranges of aircraft dynamics which are useful in aircraft design.

In 1947 Tustin reported that he had measured the frequency response of a human operator (ref. 1), an extremely novel idea at the time. Work since then has led to a 1965 report by McRuer et al (ref. 2), which includes the following:

1. A careful outline of a procedure for measuring a pilot's frequency response and noise (called a quasi-linear describing function) in a compensatory tracking task; the method relies on psychology for experimental technique and on mathematics for the Fourier analysis;
2. A large amount of supporting data including not only the pilot's frequency response (his amplitude and phase as a function of frequency), but also his remnant (the noise he injects into the control loop);
3. The crossover model of the human operator, which allows prediction of the frequency response of a pilot in a single-axis compensatory tracking task based on the vehicle dynamics.

Despite Tustin's original work in 1947 and a report by McRuer and Krendel in 1959 (ref. 3), these techniques were not immediately used due to the large quantities of data which had to be carefully processed in calculating the pilot quasi-linear-describing functions. The development of the fast Fourier transform and the hybrid computer have overcome these difficulties, and the experimental techniques and crossover model are now widely used.

State-space notation and optimal control theory are the current new tools being applied to aerospace control problems. State-space notation allows compact, systematic processing of large quantities of data, a requirement for the analysis of a pilot's tracking behavior. Such approaches have recently proven to be successful, primarily due to a simultaneous reliance upon previous work in the field of manual control, especially upon the crossover model of the human operator.

AN OPTIMAL CONTROL MODEL OF THE HUMAN OPERATOR

Preliminary History

Original attempts to use optimal control theory to model the pilot failed because, strictly speaking, a human being is not an optimal controller. For example, given the system shown in Figure 1, and knowing that only the input (i) has finite power, how can a pure gain be chosen to replace the controller to minimize mean squared error?

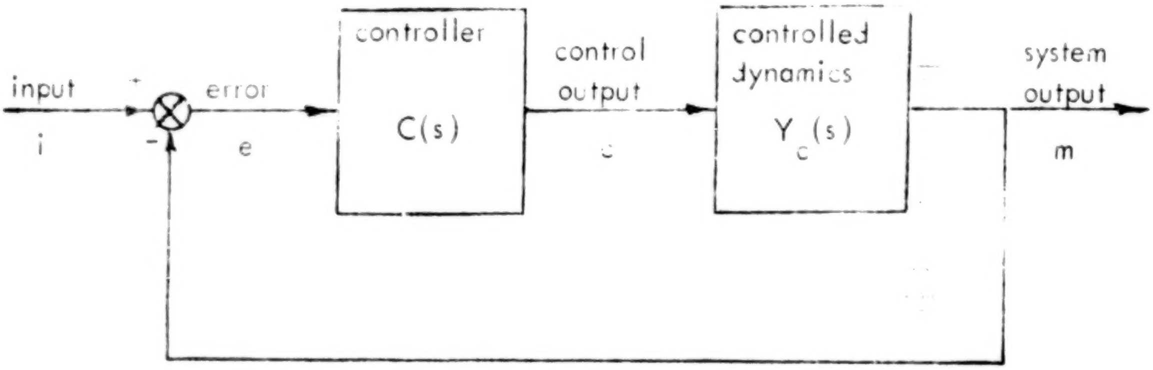


Figure 1.- A compensatory tracking system

In this simple case it is obvious that if the controller is the pure gain, K, the relationship between e and i is:

$$e(s) = [i(s)]\cdot\frac{1}{1 + Ks}$$

$$= [i(s)]\cdot\frac{s}{s + K}$$

and the mean squared error is given by

$$\overline{e(s)^2} = \overline{i(s)^2}\cdot\left|\frac{s}{s + K}\right|^2$$

Letting K go to infinity reduces the mean squared error to zero. Thus an optimal controller for this problem (in the sense of minimizing mean squared error) is an infinite gain. Unfortunately, when a pilot is placed in the controller position in Figure 1 and asked to minimize the mean squared error, he "fails": he is not an infinite gain, and the mean squared error is not zero.

TABLE OF CONTENTS

1. SPACE VEHICLE CONTROL APPLICATIONS

Design Development of the Apollo Lunar Module
Kenneth J. Cox
Manned Spacecraft Center 1 1/A10

Design Development of the Apollo Command and Service Module Thrust Vector Attitude Control System
William H. Peters
Manned Spacecraft Center 63 1/E3

Orbiting Astronomical Observatory (OAO-A2) Stabilization and Control Subsystem
Thomas E. Huber
Goddard Space Flight Center 83 1/G10

Applications of the Hybrid Coordinate Method to the TOPS Autopilot
G. E. Fleischer
Jet Propulsion Laboratory 127 2/D7

Steering Laws for Double-Gimbal Control-Moment Gyros
Stephen W. Winder
Marshall Space Flight Center 157 2/F13

2. AIRCRAFT GUIDANCE, CONTROL AND HANDLING QUALITIES

Optimum Horizon Guidance Techniques for Aircraft
Heinz Erzberger and Homer Q. Lee
Ames Research Center 175 3/A8

Inertially-Aided, All-Weather Landing Systems for C/TOL.
Robert J. Pawlak
Electronics Research Center 185 3/B4

An Approach to the Determination of Aircraft Handling Qualities Using Pilot Transfer Functions
James J. Adams and Howard G. Hatch, Jr.
Langley Research Center 199 3/C4

A Frequency Domain Approach to Handling Qualities Design
William A. Wolovich
Electronics Research Center 217 3/D8

3. SYSTEM SIMULATION AND IDENTIFICATION

Optimal Regulator or Conventional? Setup Techniques for a Model Following Simulator Control System
Dwain A. Deets
Flight Research Center 237 3/F1

TABLE OF CONTENTS (CONT'D)

3. SYSTEM SIMULATION AND IDENTIFICATION (CONT'D)

Quasi-Optimal Control of a Moving-Base Simulator
Elwood C. Stewart
Ames Research Center 253 3/G3

Some Modern Control Techniques for Human Operator Modeling and Identification
Richard S. Shirley
Electronics Research Center 261 3/G11

4. ENGINE CONTROL AND ADVANCED PROPULSION TECHNIQUES

Normal Shock and Restart Controls for a Supersonic Airbreathing Propulsion System
George H. Neiner, Gary L. Cole, and
Francis J. Paulovich
Lewis Research Center 299 4/D2

Optimal Control of a Supersonic Inlet to Minimize Frequency of Inlet Upstart
Bruce Lehtinen, John R. Zeller, and
Lucille C. Geyser
Lewis Research Center 323 4/F1

Control System Studies for Thermionic Reactors
Richard J. Hermsen,
California State Polytechnic College
Henrik G. Gronroos
JPL, California Institute of Technology 337 4/G1

Application of Fluidics to New Control Components
Miles O. Dustin, Vernon D. Gebben, and
Robert E. Wallhagen
Lewis Research Center 365 5/B7

Minimum Energy Control of a Class of Electrically Driven Vehicles
Yilmaz E. Sahinkaya
JPL, California Institute of Technology 387 5/D6

5. ADVANCED CONTROL TECHNIQUES

A New Formulation for the Epsilon Method Applied to the Minimum Time-to-Climb Problem
Lawrence W. Taylor, Jr., and Harriet J. Smith
Flight Research Center 423 5/G1

TABLE OF CONTENTS (CONT'D)

5. ADVANCED CONTROL TECHNIQUES (CONT'D)

Determination of Optimal Control Using Imbedding of the Terminal Conditions
Raymond C. Montgomery
Langley Research Center **435** 5/G13

Stability of Constant Gain Systems with Vector Feedback
George L. von Pragenau
Marshall Space Flight Center **443** 6/A10

Synthesis of Finite Settling Time Discrete Systems
James A. Gatlin
Goddard Space Flight Center **473** 6/C12

Design and Performance of Heart Assist or Artificial Heart Control Systems
John A. Webb, Jr., and Vernon D. Gebben
Lewis Research Center **495** 6/E1

At this point someone might mention that typical controls are power-limited, and that the cost functional should allow for this fact. To obtain realistic answers, the mean squared stick output ($\overline{c^2}$) must be weighted in addition to the mean squared error. If the controller is still assumed to be a pure gain, the relationship between c and i in Figure 1 is given by

$$c(s) = [i(s)]\cdot\frac{K}{1 + K/s} = [i\ (s)]\cdot\frac{Ks}{s + K}$$

and thus

$$\overline{c(s)^2} = \overline{i(s)^2}\cdot\left|\frac{Ks}{s + K}\right|^2$$

By letting K approach zero the control stick power is reduced to zero. If i(s) is defined, it is possible to calculate a value of K_J between zero and infinity which will minimize the cost functional J, given by

$$J = \overline{e^2} + k_1\overline{c^2} \tag{1}$$

where k_1 is a weighting constant. Let J_{min} be the minimum value of the cost functional. If a pilot is now placed in the position of the controller of Figure 1 and asked to minimize both mean squared error and control stick power, he still "fails": he is not a pure gain, and the value of the cost functional [Eq. (1)] is considerably larger than J_{min}. In fact, using the approach described so far (even letting the controller be other than a pure gain), it is impossible to find a cost functional which consistently leads to a controller which matches the measured behavior of the human operator (ref. 4). The reason for this is that, although the constraints of the vehicle are accounted for, the constraints of the pilot are not.

The following section outlines the development of an optimal control model of the human operator in which extensive consideration is given to his limitations. Once these additional constraints are included, it is demonstrated that, subject to his physical constraints, the well-trained human operator can be an optimal controller.

The Limitations of the Human Operator

Manual control engineers give a good indication of the limitations of the human operator which should be considered.

A large part of manual control research deals with the human operator in a single-axis compensatory tracking task of the type shown in Figure 2. A tracking task is called "compensatory" if the pilot is presented error information only, and his function is to attempt to reduce the error to zero by appropriate movements of the control stick. Compensatory tracking tasks are often found in application: examples include maintaining the position of a hovering helicopter by outside visual reference, maintaining level flight in an airplane using an artificial horizon, and following a glidepath in an airplane using a flight director. Although none of these tasks is single-axis, knowledge which is useful in analyzing multi-axis tasks has been gained from studying single-axis tasks.

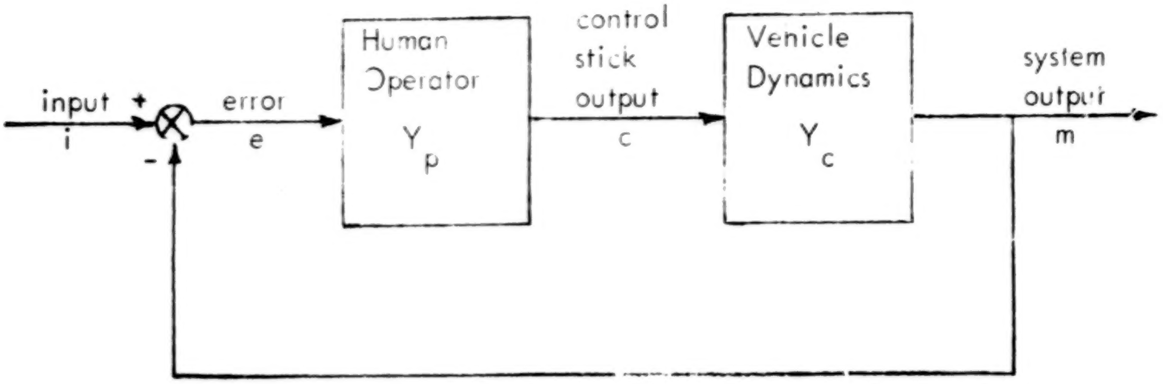


Figure 2.- The human operator in a compensatory tracking system

It has been demonstrated that the human operator can be usefully modeled by a quasi-linear-describing function as shown in Figures 3 and 4 (refs. 1-3). The quasi-linear-describing function is composed of two parts: a linear describing function, and a remnant. The linear-describing function is a transfer function, and is determined by the pilot's frequency response. In addition, by examining measured pilot frequency response curves the crossover model of the human operator has been developed (ref. 2). Specifically, it has been established that in the area of crossover* the pilot will adjust his control behavior so that the open-loop transfer function is closely approximated by

$$Y_pY_c(j\omega) = \frac{\omega_c e^{-j\omega\tau_e}}{j\omega} \tag{2}$$

---

*The crossover frequency is that frequency at which the amplitude ratio of the open-loop transfer function Y_pY_c passes from greater than unity to less than unity (see Figure 4).

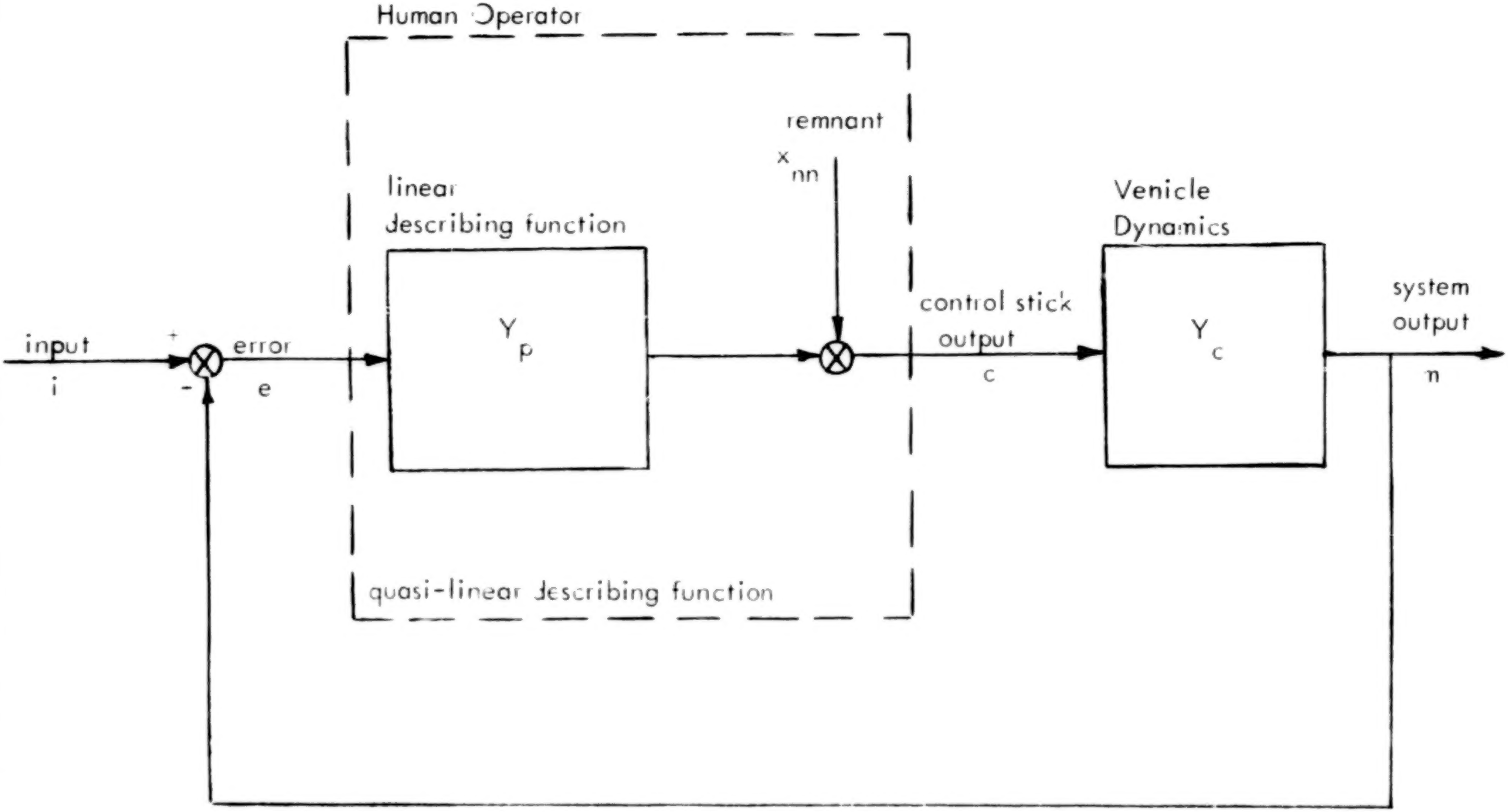

Human Operator
remnant
x_{nn}
linear
describing function
Y_p
Vehicle
Dynamics
Y_c
input
i
+
-
error
e
control stick
output
c
system
output
m
quasi-linear describing function

Figure 3.- The human operator modeled by a quasi-linear describing function

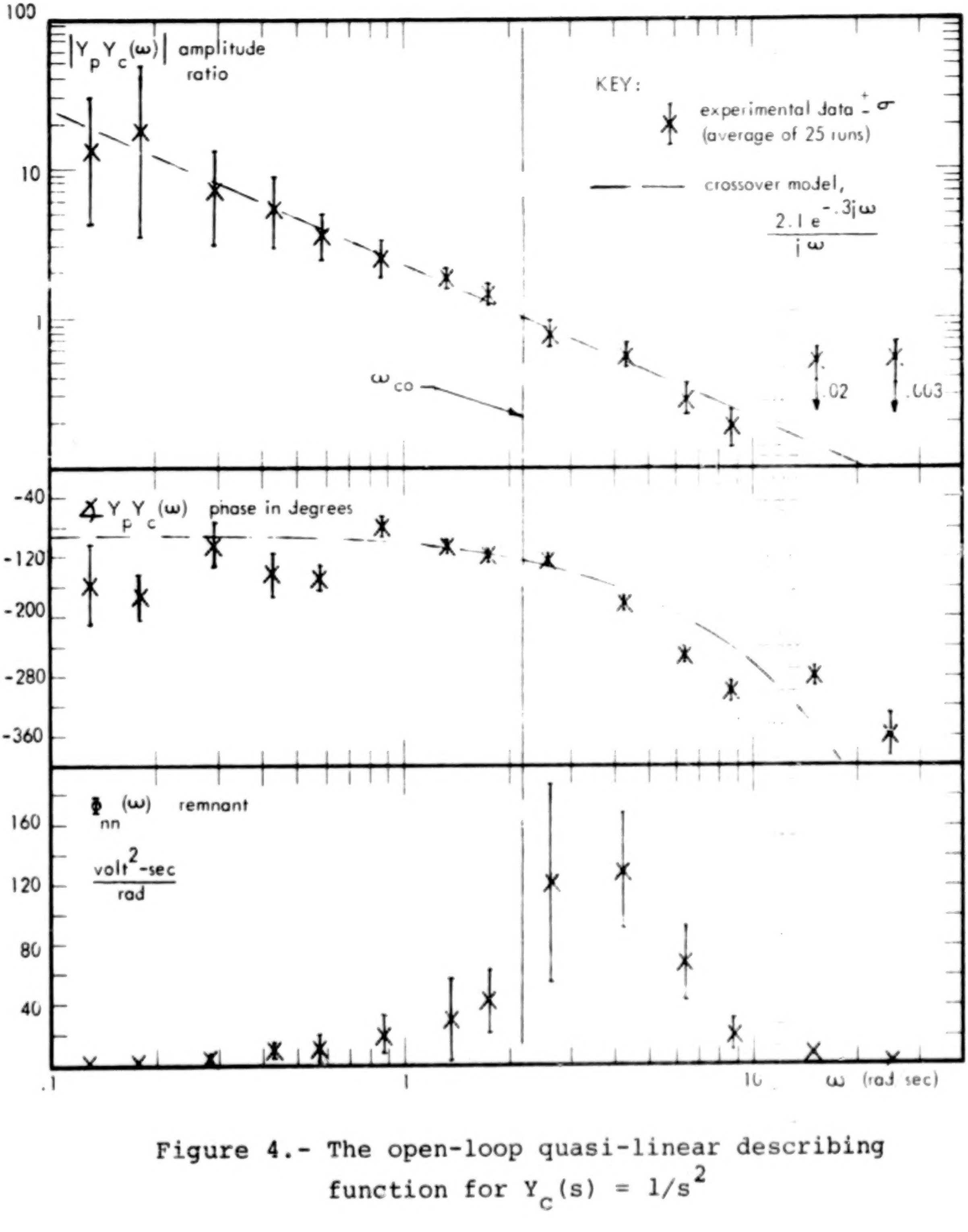


Figure 4.- The open-loop quasi-linear describing function for $Y_c(s) = 1/s^2$

Equation 2 and typical experimental data are shown in Figure 4. Refinements of this model have been made to account for the pilot's frequency response away from the crossover region. One such model is:

$$Y_p Y_c(j\omega) = K_p e^{-j\omega\tau_e} \left[\frac{T_L j\omega + 1}{T_I j\omega + 1}\right] \cdot \left[\frac{1}{T_N j\omega + 1}\right] \qquad (3)$$

Of importance in considering the pilot's physical constraints is the fact that the various terms of Eq. 3 have been associated with physical characteristics of the pilot. The exponential term is a pure time delay, and is associated with the human operator's neural processing delay of about 0.2 second, a phenomenon long known to psychologists. The lag T_N is associated with the pilot's muscular dynamics. Finally, the lead-lag pair T_L-T_I is considered to be voluntarily generated by the pilot to improve system performance.

The second portion of the quasi-linear-describing function is the remnant. The remnant is noise which the pilot injects into the man-vehicle system at the control stick. The word "remnant" is unfortunate, as it implies something left over which may be neglected. At times the linear-describing function (or the crossover model) characterizes the pilot quite well, and the remnant can be neglected. But, in the case of a hovering helicopter, for example, up to 70 percent of the control stick power is remnant, and it is unreasonable to neglect it.

In addition to the results already described, there is one other area of manual control research which is useful in developing the optimal control model of the human operator. To model the pilot on a multi-axis task, consideration is given to his instrument-scanning behavior. It has been established (refs. 5-7) that:

1. In observing a variable on an instrument a pilot obtains both position and rate information;
2. In the foveal area (direct viewing) position, information is better than rate information;
3. Information is obtained from instruments in the periphery (instruments not directly viewed);
4. In the periphery both position and rate information are degraded: rate information is slightly degraded, but position information is greatly degraded, and

5. While the eye moves from one instrument to another (a sacaad) no information is obtained by the pilot.

Thus we see that manual control engineers have often considered the following limitations of the pilot to be important: (a) instrument scanning behavior, (b) remnant (or noise), (c) neural time delay, (d) compensation capability, and (e) neuromotor dynamics. The optimal control model of the human operator was developed with these limitations in mind.

Development of the Model

Assume the human operator is in the compensatory tracking task shown in Figure 3. The diagram can be rearranged and elaborated as shown in Figure 5 (refs. 8-10). So long as the displayed variables are indications of the errors, Figure 5 still represents a compensatory tracking task. However, two very important changes have been made in going from Figure 3 to Figure 5. First, the human operator's constraints (his noise sources, neural time delays, compensation capabilities, and neuromotor dynamics) are allowed for in Figure 5. As discussed in the previous section, these parameters must be considered to develop a meaningful model of the human operator based on optimal control techniques. Second, the flow lines in Figure 5 are vector quantities. This generalization permits the model to be used to analyze and predict pilot behavior in multi-axis tasks.

In a multi-axis task more than one instrument is used to display the controlled variables. It should be noted that the vector of displayed variables y(t) includes both the variables explicitly displayed plus their first derivatives, thus including both the position and the rate information the pilot obtains while viewing the instruments. The pilot can directly view only one instrument at a time, and hence must watch the other instruments with his peripheral vision. The model thus assumes that the pilot simultaneously obtains both position and rate information from all the instruments, but with different degrees of "accuracy." The accuracy is modeled by a quantity called observation noise, shown as $\underline{v}_{\underline{y}}(t)$ in Figure 5. Each displayed variable (and its rate) thus has a noise term added when it is perceived by the pilot. The noise is assumed to be gaussian and white. The relative magnitude of the noise (its level relative to the mean squared value of the displayed variables) is a function of four parameters:

1. Whether the observed variable is displayed or is a derivative;

2. Whether the observed variable is in the foveal viewing area or in the periphery;

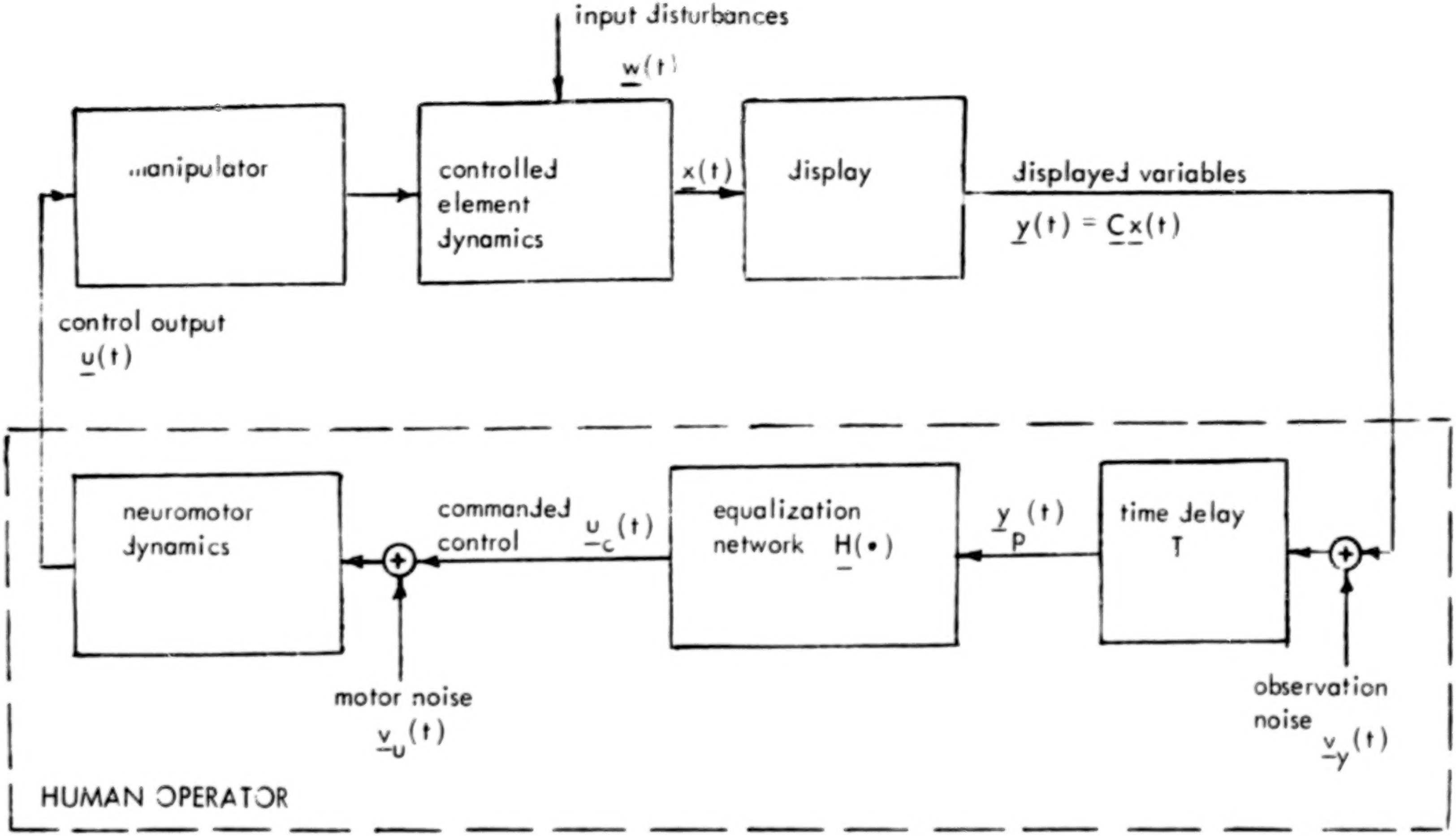

input disturbances
$\underline{w}(t)$
manipulator
controlled element dynamics
$\underline{x}(t)$
display
displayed variables
$\underline{y}(t) = \underline{\underline{C}}\,\underline{x}(t)$
control output
$\underline{u}(t)$
neuromotor dynamics
commanded control
$\underline{u}_c(t)$
equalization network $\underline{H}(\bullet)$
$\underline{y}_p(t)$
time delay T
motor noise
$\underline{v}_u(t)$
observation noise
$\underline{v}_y(t)$
HUMAN OPERATOR

Figure 5.- Model of the human operator as an optimal controller, first stage

3. The physical characteristics of each display; and

4. The relative locations of the instruments on the display panel.

Thus, the covariance matrix of the observation noise, given by

$$\underline{V}_y = E[\underline{v}_y(t)\cdot\underline{v}_y'(t)]$$

is a function not only of the display format, but also of which instrument the pilot is viewing directly. There will be as many different covariance matrices as there are instruments. It should be emphasized that the observation noise is not necessarily related only to the visual process, but can also be due to other pilot characteristics. The observation noise is thus meant to include all the noise sources in the pilot except his motor noise.

It is necessary to determine the entries in each observation noise covariance matrix in order to use the optimal control model of the human operator. Fortunately, this can be done with a series of relatively simple experiments. It is well known (refs. 2,8) that a pilot in a compensatory tracking system requires primarily position information if the dynamics are a pure gain K and primarily rate information if the dynamics are a pure acceleration K/s^2. By using these dynamics with the given display format, it is possible to measure the noise levels on both position and rate information for all the instruments as each instrument is fixated (maintained in the foveal viewing area) for a well-trained subject. These numbers are then the entries which form the covariance matrices, $\underline{V}_y$.

After the noise has been added to the displayed variables (and their first derivatives), they are delayed by the time delay of T seconds to simulate the pilot's neural time delay (both the delay mentioned in the previous section plus any other source of delay present). The result is the perceived variables, $\underline{y}_p(t)$ of Figure 5, which are a noisy, delayed version of $\underline{y}(t)$, the actual displayed variables.

The block "equalization network" of Figure 5 is shown in detail in Figure 6. The use of a Kalman filter and a predictor is based on the assumption that the well-trained human operator will attempt to reconstruct the entire current state of the vehicle, $\underline{x}(t)$, from the perception $\underline{y}_p(t)$ of the displayed variables $\underline{y}(t)$. The best estimate of the vehicle state is denoted in Figure 6 by $\hat{\underline{x}}(t)$ which is then weighted by an optimal matrix of pure gains $-\underline{l}^*$ to give a desired set of control actions $\underline{u}_c(t)$. The problem of combining the predictor with the Kalman

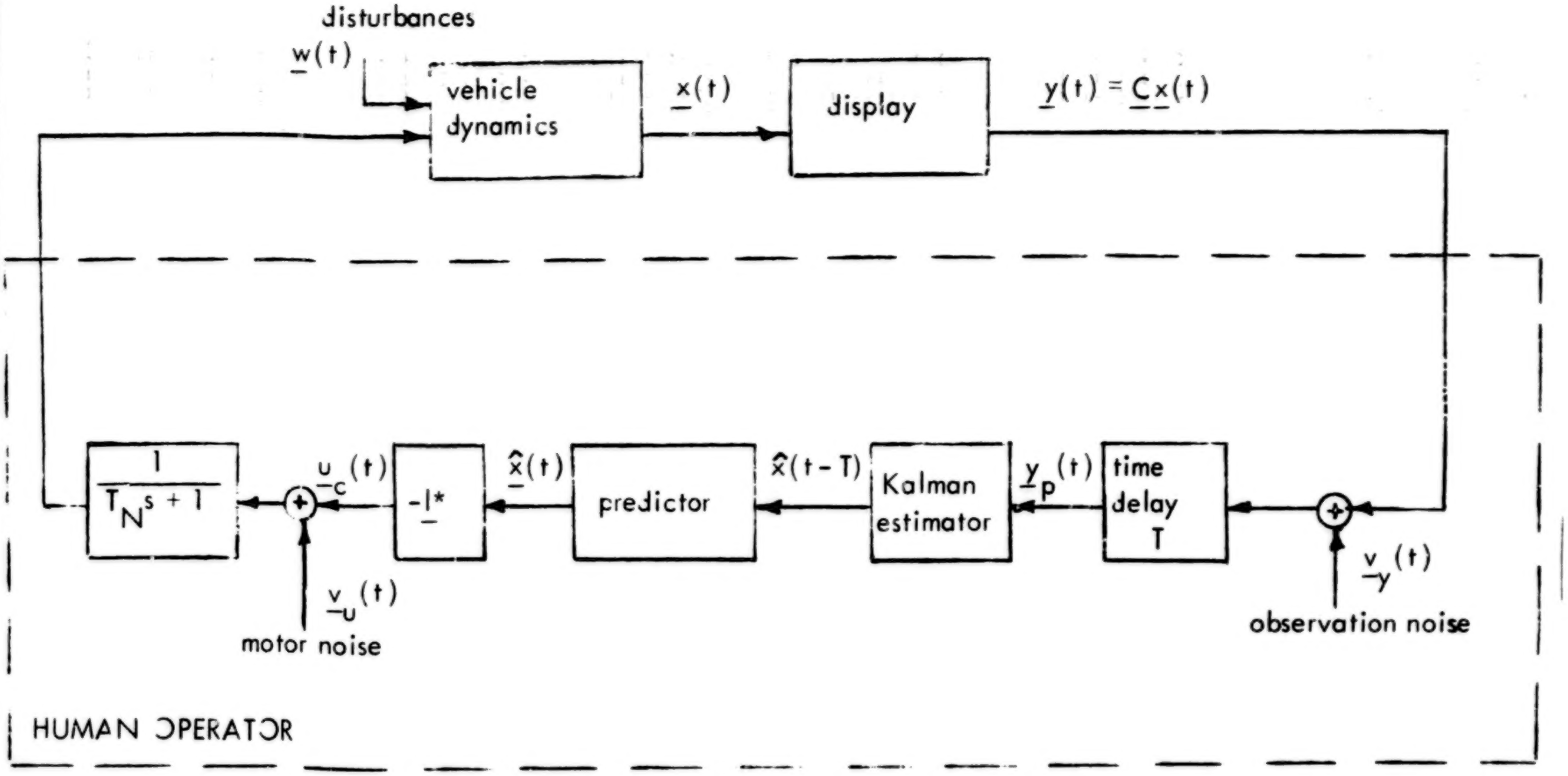

disturbances
w(t)
vehicle dynamics
x(t)
display
y(t) = Cx(t)
1 / T_N s + 1
u_c(t)
-l*
x̂(t)
predictor
x̂(t - T)
Kalman estimator
y_p(t)
time delay T
v_y(t)
observation noise
v_u(t)
motor noise
HUMAN OPERATOR

Figure 6.- Model of the human operator as an optimal controller, second stage

filter and solving the matrix-Ricatti equation to obtain the optimum gains is the contribution made to the model by optimal control theory.

The problem can now be stated as follows: the system dynamics are represented by the linear equations

$$\underline{\dot{x}}(t) = \underline{Ax}(t) + \underline{b}u(t) + \underline{w}(t)$$

where

$\underline{x}(t)$ is the vehicle state vector

$u(t)$ is the pilot's control input

$\underline{w}(t)$ is the zero-mean gaussian white noise disturbance vector, and

$\underline{A}, \underline{b}$ represent the vehicle dynamics

The displayed variables are given by

$$\underline{y}(t) = \underline{cx}(t)$$

where

$\underline{y}(t)$ is the display vector

$\underline{c}$ determines which states (and hence their first derivatives) are displayed.

As presented, $u(t)$ is a scalar. The generalization to the multi-input/multi-output case is straightforward (refs. 8, 9), but is omitted here for clarity. The human operator perceives the displayed variables as

$$\underline{y}_p(t) = \underline{c}\ \underline{x}(t-T) + \underline{v}_y(t-T)$$

(where $\underline{v}_y(t)$ is the observation noise as previously defined) and desires to generate a control, $u^*(t)$ which will minimize

$$J(u) = \sum_{i=I}^{n} q_i \overline{x_i^2} + r\ \overline{u^2} + g\ \overline{u^2}$$

The choice of this cost functional J is based on physical intuition and mathematical tractability.

The solution of this problem is given in the steps below. It is shown (refs. 8,11,12) that given a best estimate of the current state of the vehicle [$\underline{\hat{x}}(t)$], the optimal control $u^*(t)$ is generated by the linear feedback law

$$T_N\dot{u}^*(t) + u^*(t) = -\underline{l}^*\underline{\hat{x}}(t) \overset{\Delta}{=} u_c(t)$$

where $\underline{l}^*$ is determined by solution of the n+1 dimensional matrix-Ricatti equation. Furthermore, the best estimate of the state is determined using a Kalman filter and a least mean-squared predictor. This solution is implemented by the following equations which form the optimal control model of the human operator:

Step 1: Choose $\underline{Q}_o$, r, and g, where

$$\underline{Q}_o = \mathrm{diag}[q_1,q_2,\ldots,q_n,r)$$

and solve for $\underline{l}^*$ and T_N using

$$\underline{l}^* = \underline{b}_o'\underline{K}_o/g; \qquad T_N = 1/(l_{n+1}) \tag{4}$$

where

$$\underline{b}_o = [0,0,\ldots,0,1]'$$

and $\underline{K}_o$ is obtained from

$$\underline{A}_o'\underline{K}_o + \underline{K}_o\underline{A}_o + \underline{Q}_o - \underline{K}_o\underline{b}_o\underline{b}_o'\underline{K}_o/g = \underline{0} \tag{5}$$

where

$$\underline{A}_o = \left[\begin{array}{c|c} \underline{A} & \underline{b} \\ \hline \underline{0} & \underline{0} \end{array}\right]$$

This step determines T_N and $-\underline{l}^*$ as shown in Figure 6.

Step 2: Determine $\underline{V}_y$ by experiment, and solve for $\underline{\Sigma}$ using

$$\underline{0} = \underline{A}\,\underline{\Sigma} + \underline{\Sigma}\,\underline{A}' + \underline{W} - \underline{\Sigma}\,\underline{c}'\,\underline{V}_y^{-1}\,\underline{c}\,\underline{\Sigma} \qquad (6)$$

where

$$\underline{W} = E[\underline{w}(t)\cdot\underline{w}'(t)]$$

Step 3: Guess a value of T. The dynamics of the estimator are then given by

$$\dot{\underline{\hat{x}}}(t-T) = \underline{A}\,\underline{\hat{x}}(t-T) + \underline{\Sigma}\,\underline{c}'\,\underline{v}_y^{-1}\cdot[\underline{y}_p(t) - \underline{c}\,\underline{\hat{x}}(t-T)] \qquad (7)$$

$$+ \underline{b}\,u^*(t-T)$$

Step 4: The dynamics of the predictor are given by

$$\underline{\hat{x}}(t) = \underline{\xi}(t) + e^{\underline{A}t}\,[\underline{\hat{x}}(t-T) - \underline{\xi}(t-T)] \qquad (8)$$

$$\dot{\underline{\xi}}(t) = \underline{A}\,\underline{\xi}(t) + \underline{b}\,u^*(t) \qquad (9)$$

Step 5: The dynamics of the controller are given by

$$T_N\dot{u}^*(t) + u^*(t) = -\underline{1}^*\underline{x}(t) \triangleq u_c(t) \qquad (10)$$

using $\underline{1}^*$ as determined in step 1.

Step 6: The results of steps 2, 3, 4, and 5 can be combined to give the transfer function of the pilot-vehicle combination as

$$\underline{H}(s) = -\frac{[\underline{1}^*,0\,]}{T_N s + 1}\left\{(s\underline{I} - \underline{A})\int_0^T e^{(s\underline{I} - \underline{A}_1)\sigma}\,d\sigma\cdot(s\underline{I} - \bar{\underline{A}}) \qquad (11)\right.$$

$$\left. + s\underline{I} - \hat{\underline{A}} + \underline{b}_1[\underline{1}^*,0]\right\}^{-1}\cdot\underline{\Sigma}\,\underline{c}'\,\underline{V}_y^{-1}$$

where

$$\hat{\underline{A}} = \underline{A}_1 - \underline{\Sigma}\, \underline{c}_1\, \underline{V}_y^{-1}\, \underline{c}_1$$

$$\bar{\underline{A}} = \left|\begin{array}{c:c} \underline{A} & \underline{b} \\ \hdashline \underline{1}^*/T_N & -1/T_N \end{array}\right|$$

$$\underline{A}_1 = \left|\begin{array}{c:c} \underline{A} & \underline{b} \\ \hdashline \underline{0} & -1/T_N \end{array}\right|$$

$$\underline{b}_1 = \mathrm{col}[0,0,\ldots,0,1/T_N]$$

and

$$\underline{c}_1 = [\underline{c}:\underline{0}].$$

A complete discussion and derivation of steps 1 through 6 can be found in references 8 through 14. The solution of Eqs. (4) through (11) allows determination of the human's optimal control input $u_c(t)$ as shown in Figure 6. In the absence of motor noise, the desired input passes through the neuromotor dynamics and feeds back to the vehicle. It should be noted that the value of T_N is determined by g in Eqs. (4) through (11) and hence T_N is not a free parameter. In simple tasks, if g is chosen to yield a value of T_N close to that expected for the pilot's neuromuscular time constant, a good fit to the data is obtained. T_N can include other neural effects such as a subjective rating on the stick rate, and is not necessarily related only to the muscular characteristics of the pilot.

The inclusion of motor noise $\underline{v}_u(t)$ is complex and not completely understood. A discussion of the effects of $\underline{v}_u(t)$ is found in reference 14.

The Scanning Problem

In the solution of Eqs. (4) through (11) the observation noise covariance matrix $\underline{V}_y$ was assumed constant. In the multi-axis case, however, the entries in $\underline{V}_y$ change depending upon which instrument the pilot is directly viewing. Let ω_s symbolize the pilot's sampling strategy, or instrument scanning behavior, and denote the dependence of $\underline{V}_y$ on ω_s by writing $\underline{V}_y(\omega_s)$

Note that $\underline{\Sigma}$ [Eq. (6)] is a measure of the estimation error of the Kalman filter, and is also a function of $\underline{V}_y$ and hence of ω_s. Furthermore the matrix

$$\underline{F} = \underline{1}^{*\prime} \; \underline{\Sigma} \; \underline{1}^*$$

is a measure of the estimation error weighted according to the importance of the various states as determined by the q_i. It is thus necessary to choose a sampling strategy, ω_s, to minimize $\underline{F}$. Then $\underline{V}_y(\omega_s)$ and $\underline{\Sigma}(\omega_s)$ can be used in solving Eqs. (4) to (11). Further discussion of the problem is presented in references 8, 10, 14, and 17. The solution of the problem is the predicted sampling behavior of the pilot.

Summary of the Model

Equations (4) to (11) and the optimal sampling program can currently be solved on a time-shared digital computer. The program is written to accept the variables and print out the results shown in Table I.

Evaluation of the Model

The model has been used to predict pilot performance in a series of single-axis and in one two-axis task. Figures 7, 8, and 9 (from reference 14) show model predictions and experimental measurements of pilot performance in single-axis compensatory tracking tasks with three different vehicle dynamics. The only model parameters changed in predicting pilot performance for these three figures (see Table II) are T_N and T. All other parameters (except the $\underline{A}$ matrix which specifies the change of vehicle dynamics) are held constant, yet the model quite accurately predicts average pilot performance, including remnant, for the three single-axis cases studied.

The model was also used to predict pilot performance in a two-axis tracking task. The task was to minimize the mean squared pitch and longitudinal position errors of a simulated XV-5A VTOL aircraft. Displays of both pitch and position were provided. Position was maintained by pitching the aircraft. Further details of the experiment are available in reference 9. Of interest here are the excellent predictions of the experimental data the model provided, as shown in Figure 10. It should be noted that for this experiment there were three free parameters of the model: T_N, T, and q_θ.

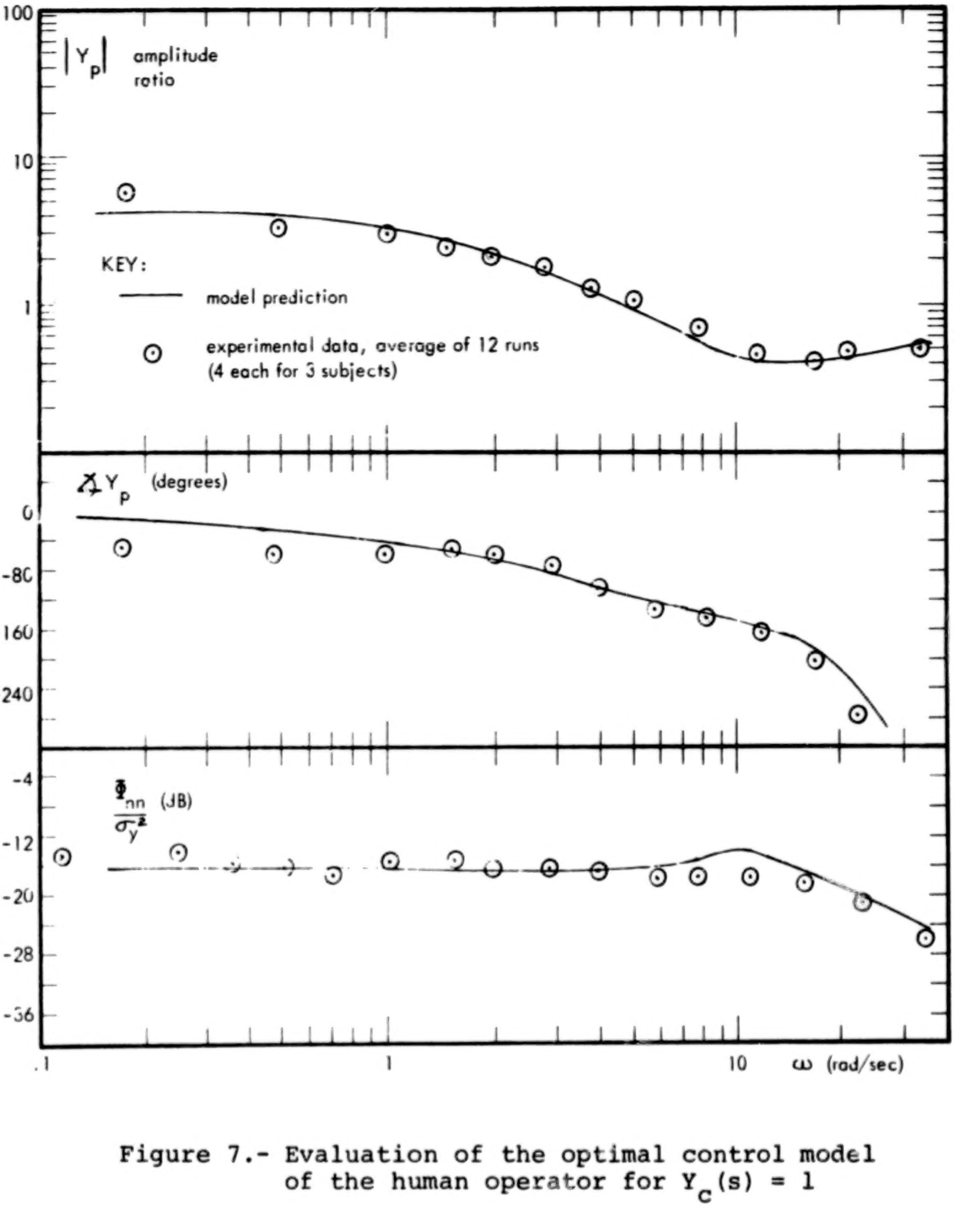


Figure 7.- Evaluation of the optimal control model of the human operator for $Y_c(s) = 1$

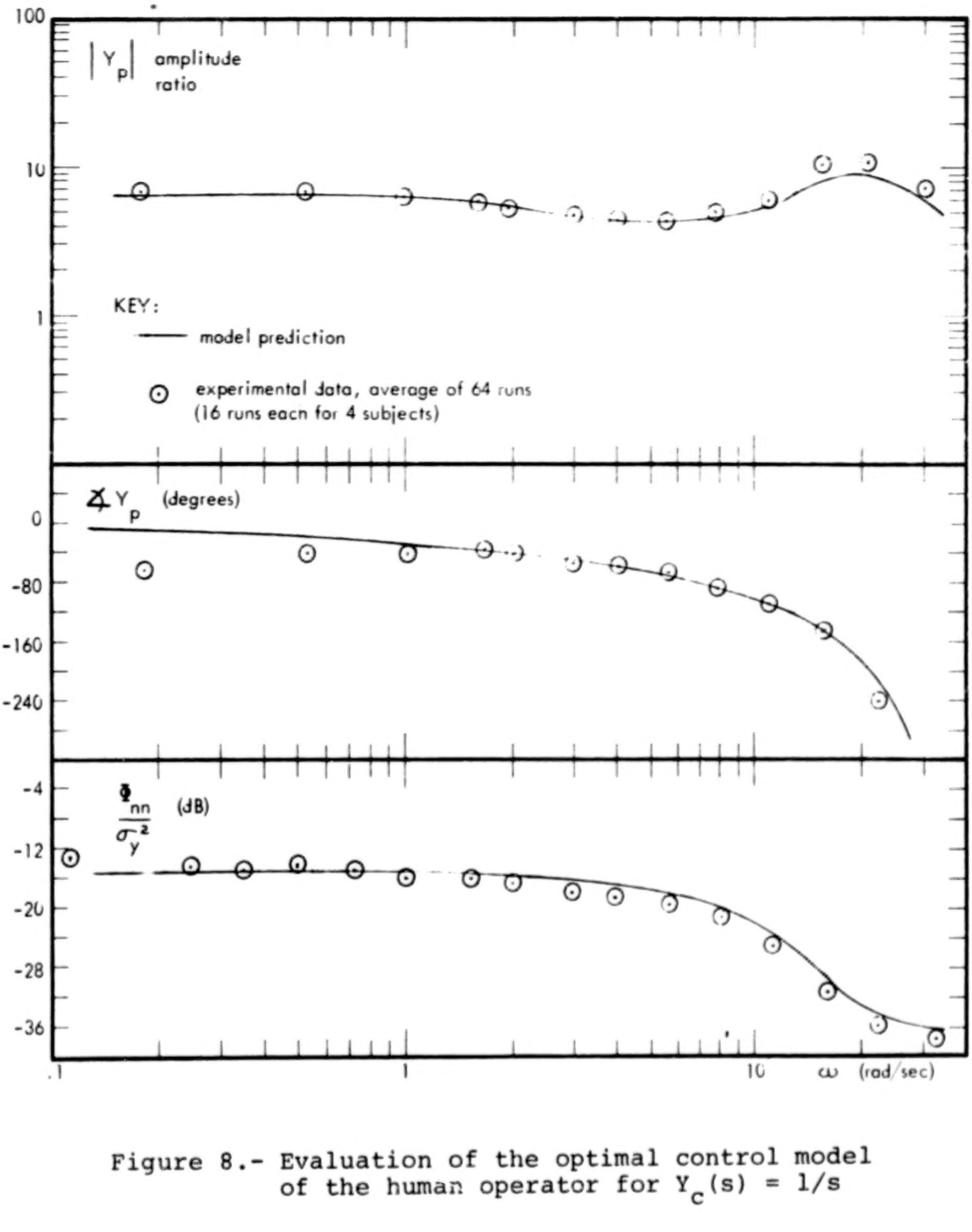


Figure 8.- Evaluation of the optimal control model of the human operator for $Y_c(s) = 1/s$

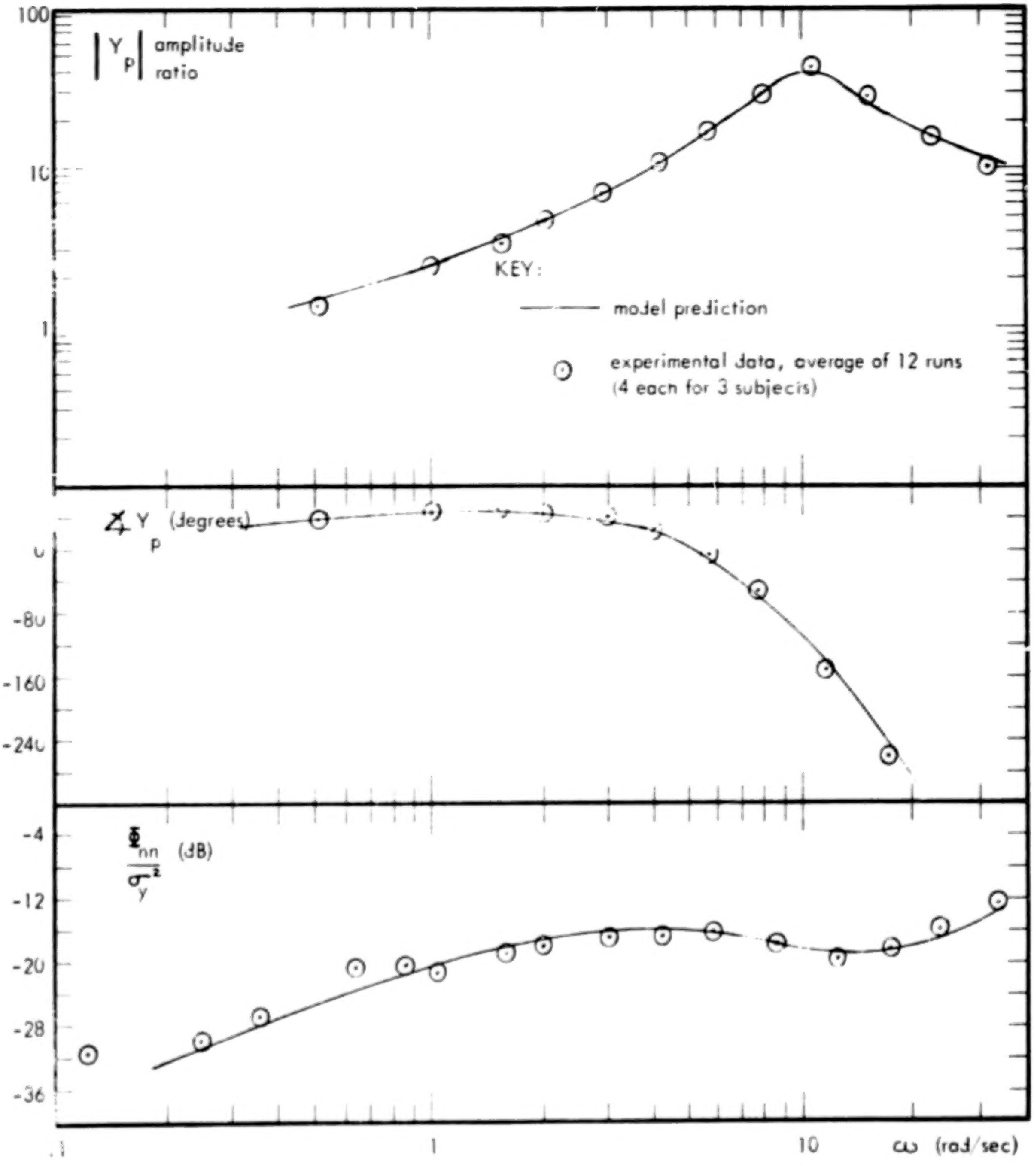


Figure 9.- Evaluation of the optimal control model of the human operator for $Y_c(s) = 1/s^2$

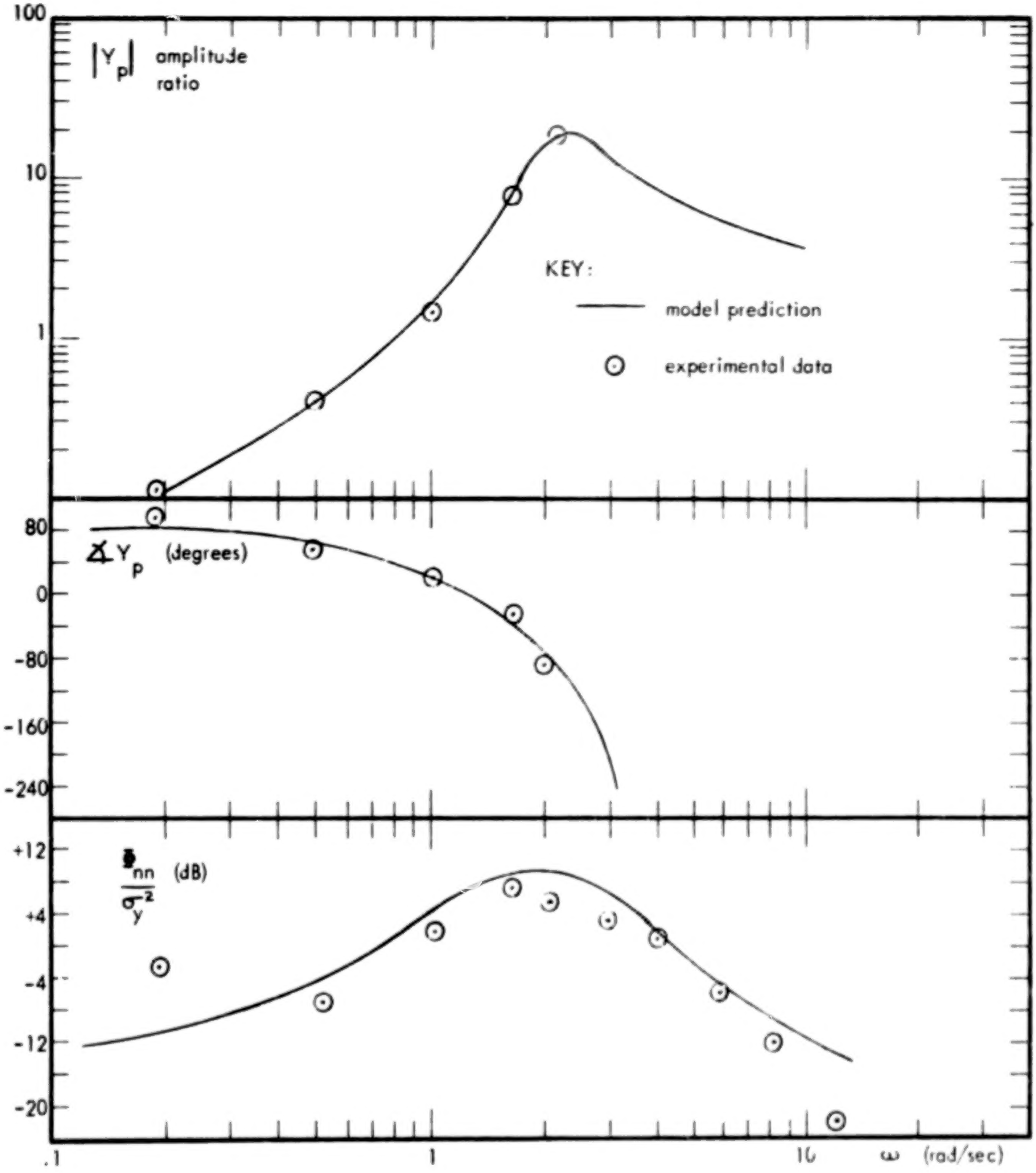


Figure 10.- Evaluation of the optimal control model of the human operator for Y_c = (dynamics of the XV-5A), longitudinal axis

TABLE I

INPUTS AND OUTPUTS FOR THE OPTIMAL CONTROL MODEL OF THE HUMAN OPERATOR

PROGRAM INPUTS

- A. Inputs determined by vehicle and disturbance
    - 1. Determined by vehicle
        - $\underline{A}$ vehicle dynamics
        - $\underline{b}$ control stick effect
    - 2. Determined by input disturbance
        - $\underline{w}$ gaussian white noise
    - 3. Determined by display
        - $\underline{c}$ choice of states displayed
        - $\underline{v}_y$ observation noise
- B. Inputs determined by human operator
    - 1. Determined by physical constraints
        - $\underline{v}_u$ motor noise
        - T neural time delay
        - T_N neuromuscular time constant (not free, see g_j below)
    - 2. Determined (mostly) by choice
        - q_i weightings on vehicle state
        - r_j weightings on control powers
        - g_j weightings on control rates (determines T_N)

PROGRAM OUTPUTS

- A. Pilot linear describing function for each axis
    - 1. Amplitude ratio versus frequency
    - 2. Phase versus frequency
- B. Pilot stick spectra
    - 1. Remnant (uncorrelated spectrum)
    - 2. Correlated spectrum
- C. Pilot instrument scanning strategy
- D. Loop error and power scores for each axis

TABLE II

VALUES OF FREE PARAMETERS USED IN TESTING MODEL
(see Figures 7, 8, and 9)

Model Parameter	Dynamics		
	K	K/s	K/s^2
T_N	0.11	0.08	0.10
T	0.15	0.15	0.21

T_N, neuromuscular time constant, set at 0.22 by choice of g

T, neural time lag, set at 0.25

q_θ, weighting on pitch, set at 2.0

(Above values apply to figure 10 for XV-5A)

APPLICATION OF A MODIFIED B.L. HO ALGORITHM TO PROCESSING PILOT DATA

Overview

The optimal control model of the human operator is designed to predict pilot performance. In order to evaluate the model, as well as for many other applications, it is necessary to measure pilot performance (his frequency response and remnant). In addition, it is often desirable to obtain a transfer function which closely fits the measured frequency response. There are four commonly used techniques for calculating the frequency response of the human operator from his input-output data. These are Fourier analysis, cross correlation, filtering, and parameter tracking. Of these four only parameter tracking directly provides a transfer function of the human operator. The other techniques require a fit to the frequency response. As the best linear model form (number of zeroes and whether poles are complex) is not always known, it is desirable to be able to obtain the transfer function without specifying the model form. A modification of an algorithm developed in 1965 by B.L. Ho provides one way for this to be done (ref. 15).

What the Algorithm of B.L. Ho Does

Given a sequence of N+1 numbers $\{z_k\}_0^N$, the B.L. Ho algorithm calculates the minimum-order matrix triplet, $\underline{c}$, $\underline{A}$, $\underline{b}$, such that

$$z_k = \underline{c}\,\underline{A}^k\,\underline{b} \tag{12}$$

for all integer k running between 0 and N.

The usefulness of the algorithm (for linear, noise-free systems) is shown as follows. A linear system can be written in the form:

$$\begin{aligned} \underline{\dot{x}} &= \underline{A}\,\underline{x} + \underline{b}\,\underline{u} \\ \underline{y} &= \underline{c}\,\underline{x} \end{aligned} \tag{13}$$

The linear system represented by Eq. (13) is said to be realized by the matrix triplet $\underline{c}$, $\underline{A}$, $\underline{b}$. From these matrices it is relatively easy to derive the transfer function between $\underline{y}$ and $\underline{u}$ if it is desired. The reverse, however, is not true: one cannot always specify a system realization from the transfer function. For zero initial conditions the solution to Eq. (13) is given by

$$\underline{x}(t) = \int_0^t e^{(t-\tau)\underline{A}}\,\underline{b}\,\underline{u}(\tau)\,d\tau \tag{14}$$

From Eq. (14) it can be seen that the impulse response is given by

$$\underline{h}(t) = \underline{c}\,e^{\underline{A}t}\,\underline{b} \tag{15}$$

In the discrete case (if h(t) is sampled):

$$\begin{aligned} \underline{h}(k\delta) &= \underline{c}\,e^{\underline{A}k\delta}\,\underline{b} \\ &= \underline{c}\,(e^{\underline{A}\delta})^k\,\underline{b} \end{aligned} \tag{16}$$

or, defining $\underline{F} = e^{\underline{A}\delta}$

$$\underline{h}\,(k\delta) = \underline{c}\,\underline{F}^k\,\underline{b} \tag{17}$$

Equations (12) and (17) show that if the series $\{z_k\}_0^N$ is the sampled impulse response $\underline{h}\,(k\delta)$, the method of B. L. Ho allows direct calculation of $\underline{C}$, $\underline{F}$, and $\underline{b}$. It is then an easy matter to use a matrix log program to derive $\underline{A}$ from $\underline{F}$, and hence the matrix coefficients of Eq. (13) are determined: a minimum order system realization has been obtained.

There is a second way in which B. L. Ho's algorithm can be used to obtain a linear system realization. Denote the k^{th} derivative of the system impulse response at zero time by $\underline{h}^{(k)}(0)$. From Eq. (15) it is obvious that

$$\underline{h}^{(0)}(0) = \underline{c}\ \underline{b}$$

$$\underline{h}^{(1)}(0) = \underline{c}\ \underline{A}\ \underline{b} \tag{18}$$

$$\underline{h}^{(k)}(0) = \underline{c}\ \underline{A}^k\ \underline{b}$$

Comparison of Eqs. (18) and (12) shows that if the series $\{z_k\}_o^N$ is composed of the successive derivatives of the impulse response at zero time, B. L. Ho's algorithm will yield a minimum order linear system realization. There is, however, the problem of obtaining the required derivatives. It is certainly impractical to attempt to calculate them directly from the impulse response. Fortunately, it is known that the expansion of the transfer function in terms of the Laplace variable s can be written as

$$\underline{H}(s) = \underline{h}^{(0)}(0)\ s^{-1} + \underline{h}^{(1)}(0)\ s^{-2} + \ldots + \underline{h}^{(k)}(0)\ s^{-(k+1)} \tag{19}$$

where the coefficients of $s^{-(k+1)}$ are the desired derivatives.

Thus the B. L. Ho algorithm can be used to obtain a minimum order linear system realization from either the noise-free (exact) system impulse response, or from the system transfer function, as summarized on Figure 11.

The B. L. Ho Algorithm

Given the sequence $\{z_k\}_o^N$, assume

$$z_k = \underline{c}\ \underline{A}^k\ \underline{b}$$

then

$$\underline{A} = \underline{P}_r\ \underline{\bar{H}}_d\ \underline{Q}_r \tag{20}$$

$$\underline{b} = \underline{P}_r\ \underline{\xi} \tag{21}$$

$$\underline{c} = \underline{\xi}'\ \underline{Q}_r \tag{22}$$

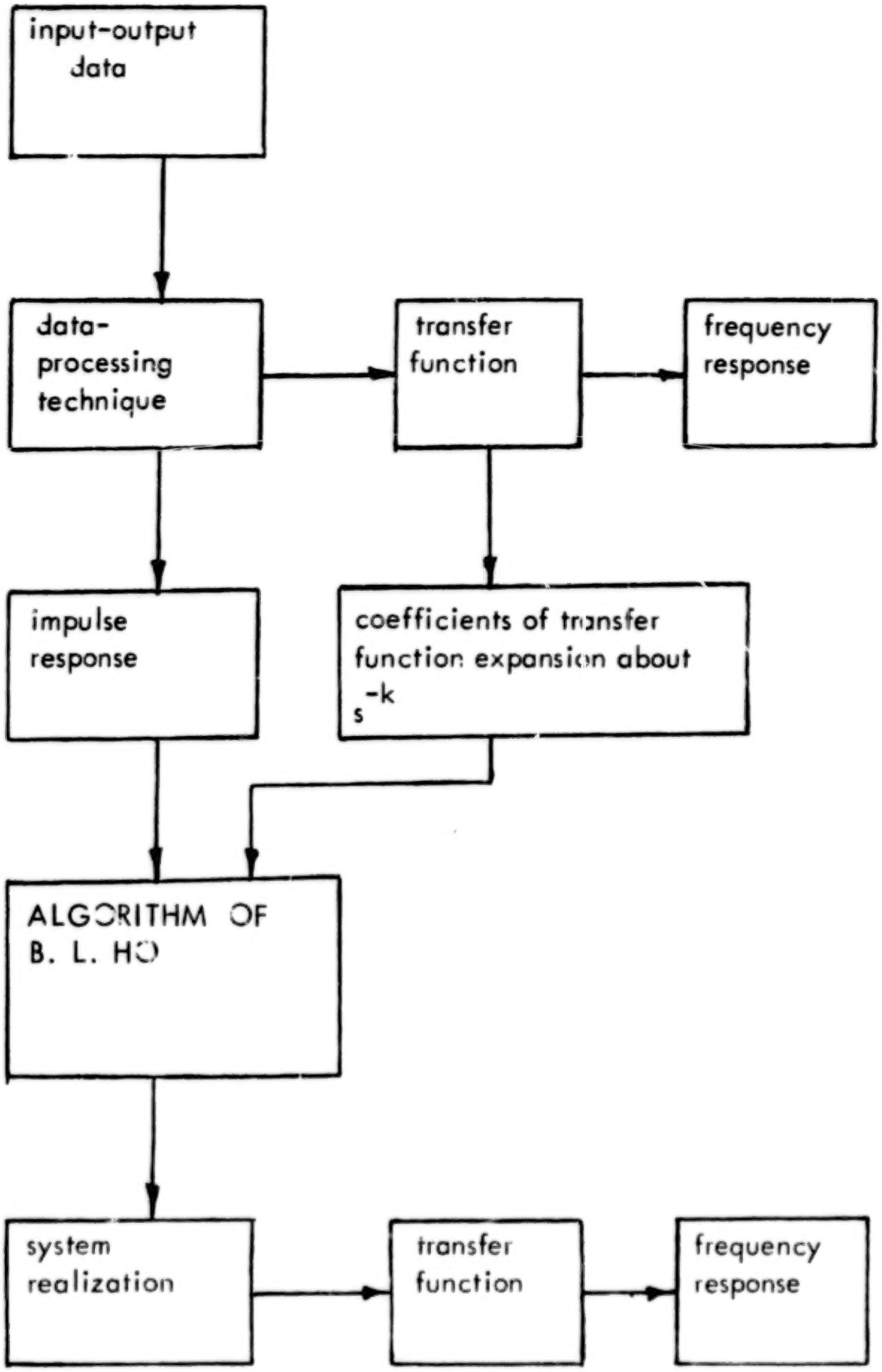

input-output data
data-processing technique
transfer function
frequency response
impulse response
coefficients of transfer function expansion about s^{-k}
ALGORITHM OF B. L. HO
system realization
transfer function
frequency response

Figure 11.- What the B.L. Ho algorithm can be used for

Equations (20), (21), and (22) provide the minimum order linear system realization $\underline{c}$, $\underline{A}$, $\underline{b}$. The other matrices in these equations are defined as follows:

$$\underline{\xi} = \mathrm{col}[z_0, z_1, z_2, \ldots, z_d] \tag{23}$$

$$\underline{\overline{H}}_d = \begin{bmatrix} z_1 & z_2 & z_3 & \cdots & z_d \\ z_2 & z_3 & & & \vdots \\ z_3 & & & & \\ \vdots & & \ddots & & \\ z_d & \cdots & & & z_{2d-1} \end{bmatrix} \tag{24}$$

(assuming $2d<N$)

$\underline{P}_r$ = first r rows of $\underline{P}_d$

$\underline{Q}_r$ = first r columns of $\underline{Q}_d$

and $\underline{P}_d$ and $\underline{Q}_d$ are any matrices such that

$$\underline{P}_d \, \underline{H}_d \, \underline{Q}_d = \begin{bmatrix} \underline{I}_r & 0 \\ 0 & 0 \end{bmatrix} \tag{25}$$

where I_r is the rth order identity matrix and

$$\underline{H}_d = \begin{bmatrix} z_0 & z_1 & z_2 & \cdots & z_{d-1} \\ z_1 & z_2 & z_3 & & \vdots \\ z_2 & & & \ddots & \\ \vdots & & & & \\ z_{d-1} & \cdots & & & z_{2d-2} \end{bmatrix} \tag{26}$$

The problem of finding a $\underline{P}_d$ and a $\underline{Q}_d$ given $\underline{H}_d$ such that Eq. (25) holds is a standard computational technique which is part of most generalized inverse routines. In practice, one starts with d=2 and finds a $\underline{P}_2$ and $\underline{Q}_2$ such that

$$\underline{P}_2\ \underline{H}_2\ \underline{Q}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Then d is incremented and the process repeated until Eq. (25) is satisfied.

Equations (20) through (26) form the B. L. Ho algorithm. The proof of the algorithm can be found in references 16 and 17.

Application to Processing Human Operator Data

The method of B. L. Ho assumes that the system realized by the matrix triplet $\underline{c}$, $\underline{A}$, $\underline{b}$, is linear, and that the series $\{z_k\}_0^N$ is free from the effects of noise (that the impulse response is exact). Strictly speaking, neither of these assumptions is valid in the case of the human operator. If it is assumed that the system being analyzed is the linear portion of the quasi-linear describing function as described earlier in this paper, the assumption of linearity is somewhat justified. The noise-free assumption is not justified, however, and the effects of noise upon the system realization obtained must be considered and allowed for.

In the current application three steps were taken to reduce the effects of noise. First, an attempt was made to reduce the errors in the impulse response due to noise by processing the data from relatively long runs (216 seconds). Secondly, the algorithm of B. L. Ho was modified by extending the row dimension of $\underline{H}_d$ and $\underline{\bar{H}}_d$ in Eqs. (20) through (26), plus making appropriate changes to the other matrices to maintain conformability. The increased row dimensions of $\underline{H}_d$ and $\underline{\bar{H}}_d$ permit N samples of the impulse response to influence the system realization instead of the first 2d samples only (ref. 18). Finally, in implementing the modified B. L. Ho algorithm on a digital computer, one never really gets a zero in the lower right-hand corner of the matrix $\underline{P}_d\underline{H}_d\underline{Q}_d$ [Eq. (25)]. A tolerance is chosen, and when the lower right-hand corner of Eq. (25) falls below this tolerance (called Tol. 1), it is considered zero and the rank is determined as r. There is thus a tradeoff between Tol. 1 and the system rank which must be experimentally determined.

The Experiment

Data were obtained from a compensatory tracking system at i, e, c, and m as shown in Figure 2. The experiment was run using the sum-of-sinusoids forcing function shown in Table III. As the data were being stored on magnetic tape they were also processed using the method of Fourier coefficients as described in reference 19. The results of the Fourier analysis are used to check the results obtained by: (1) using a filtering technique to obtain an impulse response, (2) the modified B. L. Ho algorithm to obtain a system realization, and (3) calculation of the frequency response from the system realization The last step was made only to permit direct comparison of frequency response curves.

Four experiments were performed as summarized in Table IV. In all cases the impulse response was obtained using a filtering technique. The filters were a series of poles in the s-plane as shown in Table V. On a digital computer the output of each filter in response to the error signal e was calculated. A least squares technique was then used to weight the filter outputs so as to minimize the difference between the actual system output m and the sum of the weighted filter outputs. These filter coefficients were then used to weight the known filter impulse responses which were added together to approximate the impulse response of the system being identified. In the process, steps were taken to effectively eliminate the effects of non-zero initial conditions. The noise effects were reduced, but not eliminated, by averaging over the 216-second run time. Thus the calculated impulse responses used as input to the modified B. L. Ho algorithm were inexact due to the human operator's remnant.

TABLE V

THE FILTERS USED TO OBTAIN THE SYSTEM IMPULSE RESPONSE

Real poles at:	Complex Poles at:
$s = -.269$	$s = -.722 \pm .25j$
$s = -.520$	$s = -1.39 \pm .45j$
$s = -1.00$	$s = -2.68 \pm .85j$
$s = -1.93$	
$s = -3.70$	

TABLE III

PARAMETER VALUES FOR THE PSUEDO-RANDOM FORCING FUNCTION

k	A_k (volts)	ω_k (rad/sec)
1	.2	26.18
2	-.2	15.71
3	.2	8.727
4	-.2	6.545
5	.2	4.363
6	-1.	2.618
7	1.	1.745
8	-1.	1.309
9	1.	.8727
10	-1.	.5818
11	1.	.4363
12	-1.	.2909
13	1.	.1745
14	-1.	.1164

$i(n\Delta t) = A_k \sin(\omega_k n\Delta t)$

$\Delta t = 0.02$ second

T = data-taking period = 216 seconds

N = number of data samples taken = 10,800

TABLE IV

SUMMARY OF EXPERIMENTS

Experiment Number	Y_p	Y_c	Noise	Dynamics Identified	Signals Processed Input, Output
1	$1/(s+1)$	$1/(s+1)$	none	$Y_p Y_c$	e,m
2	$1/(s+1)$	$1/(s+1)$	white noise at m	$Y_p Y_c$	e,m
3	human subject	$1/s$	generated by pilot	Y_p	e,c
4	human subject	$1/s^2$	generated by pilot	Y_p	e,c

In the first experiment only (see Table IV) there was no noise, so any errors in the system realization were due to the experimental technique. The realization obtained had poles at -0.99 and -1.01, a close approximation to the actual system with two poles at -1. The frequency response of this realization, the calculated frequency response of the known filter, and the frequency response measured by the method of Fourier coefficients are shown on Figure 12.

When white noise was added to the system output, however, the impulse response obtained from the data was not exact. The errors in the impulse response caused the modified B. L. Ho algorithm to produce a system of maximum allowed order (specified by the programmer) or a fourth-order system in this case. Furthermore, reduction of the tolerance value mentioned previously did not reduce the order of the realization until the tolerance was so large as to lead to a poor fit (see Figure 13). This relationship held true in the case of data from human operators, as well. Hence it is necessary to specify the maximum allowable system order in using the modified B. L. Ho algorithm, unless the linear system impulse response is almost exactly determined (as in Figure 12). The form of the model (that is, the number of zeroes and the number of zeroes or poles which are complex) is determined by the modified algorithm, and is not specified by the programmer.

Figures 14 and 15 show the frequency response of fourth-order linear realizations of the human operator in a compensatory tracking system controlling the dynamics 1/s and $1/s^2$, respectively. If we allow for the fact that these fits are based on a single 216-second run, they compare well with results in the literature (refs. 2, 3) in the critical crossover region (near ω_{co}).

CONCLUSION

Some modern control techniques have been successfully applied to modeling and identifying the human operator in a manual control task. Results reviewed in this paper include:

1. A technique for predicting the pilot's remnant and error scores in addition to his linear describing functions in a multi-axis task;

2. A method for obtaining a linear system realization of the human operator from his input-output data via an impulse response without assuming model form (the number of zeroes or the number of poles which are complex).

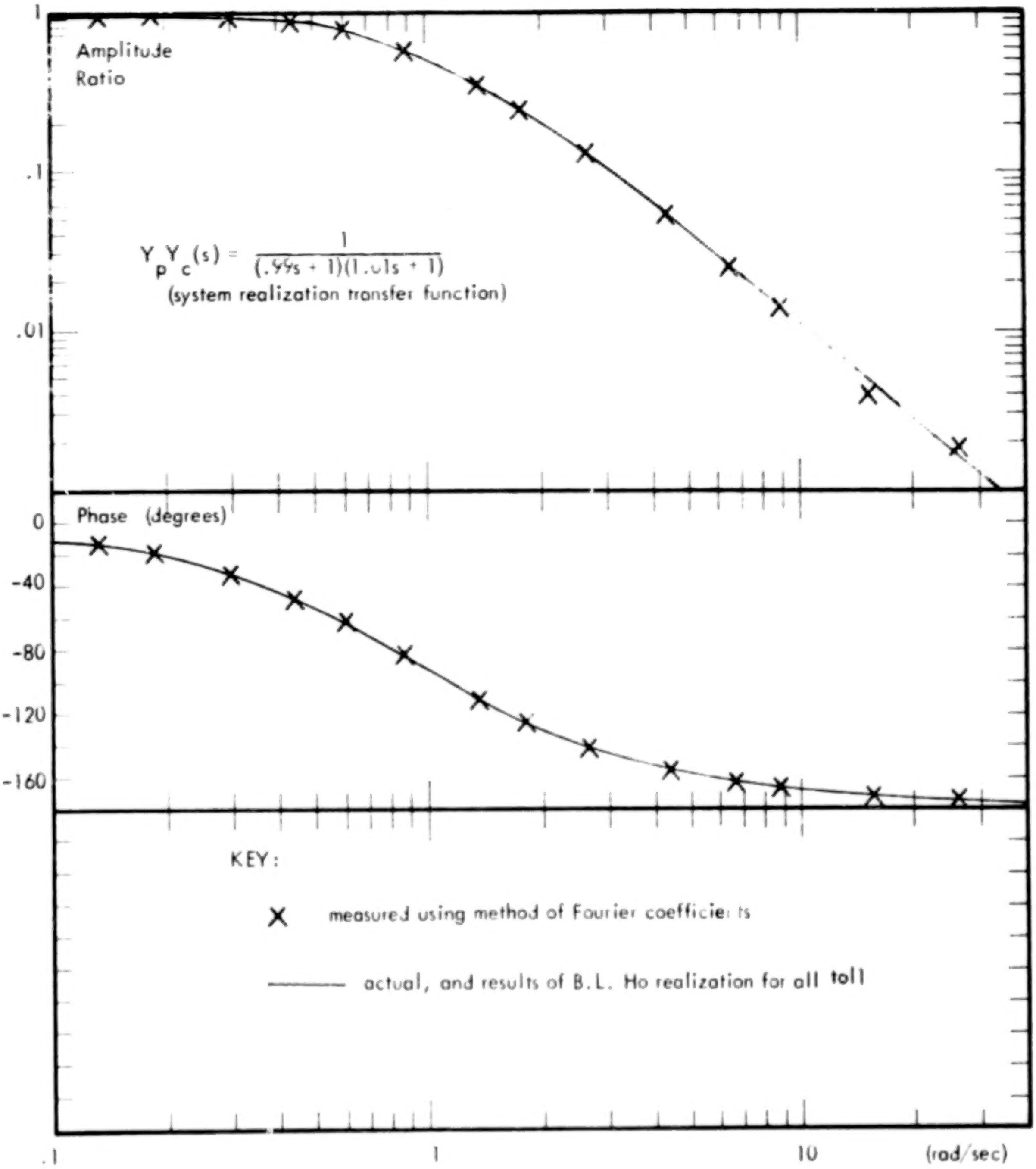


Figure 12.- Frequency response of B.L. Ho system realization for $1/(s+1)^2$, no noise

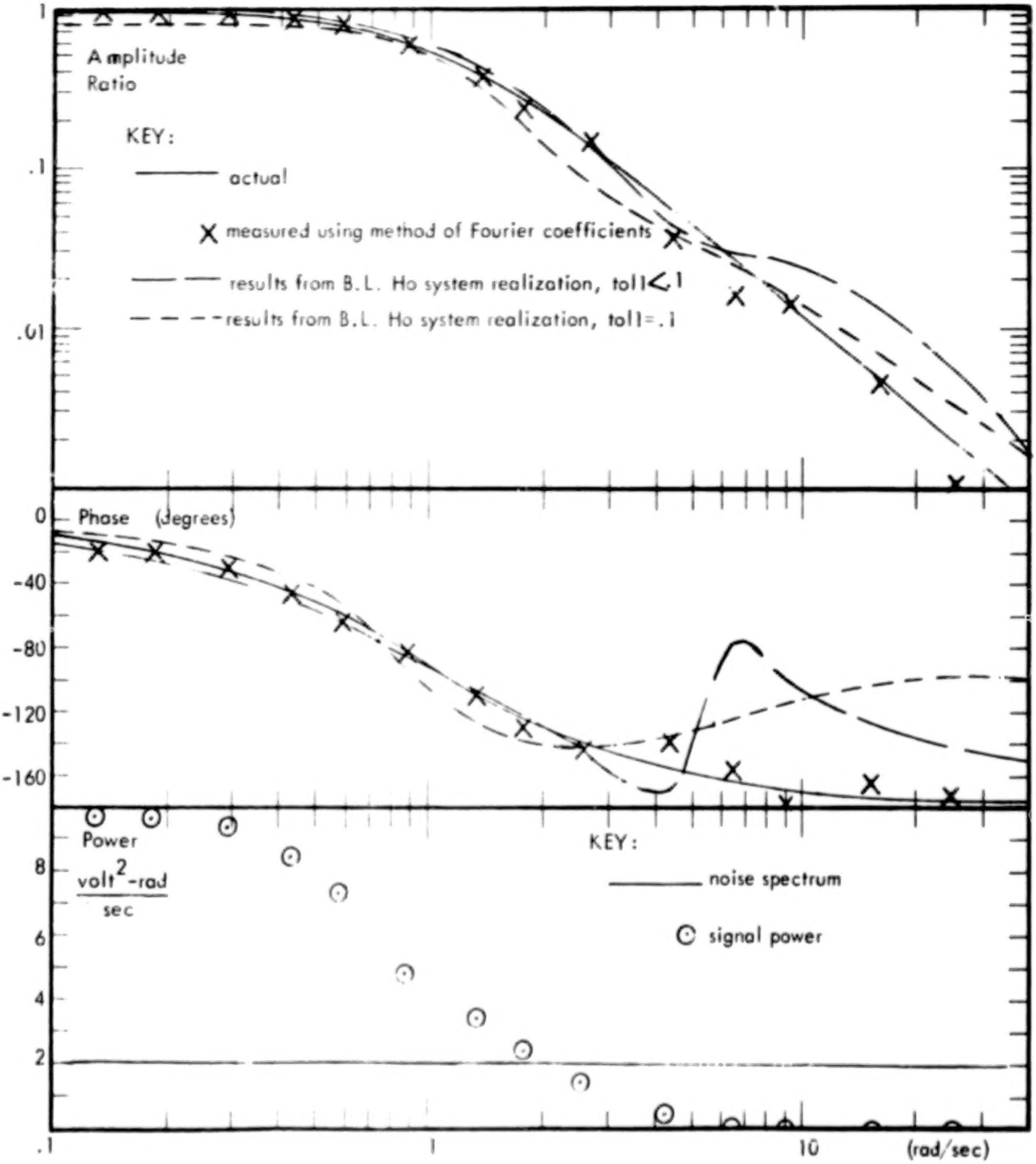


Figure 13.- Frequency response of B.L. Ho system realization for $1/(s+1)^2$, with noise, S/N = 1.0

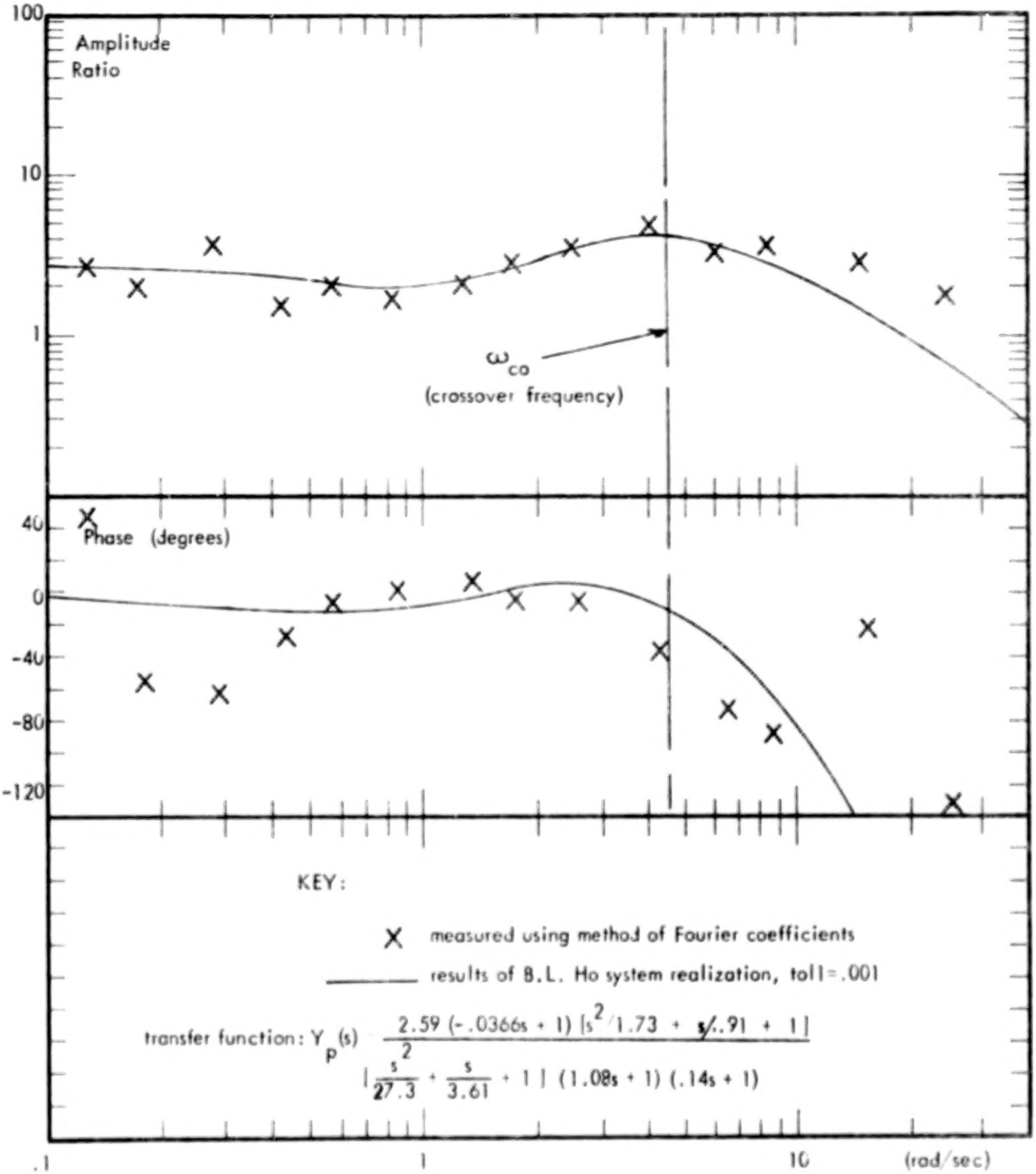


Figure 14.- Frequency response of B.L. Ho system realization for human operator, $Y_c(s) = 1/s$

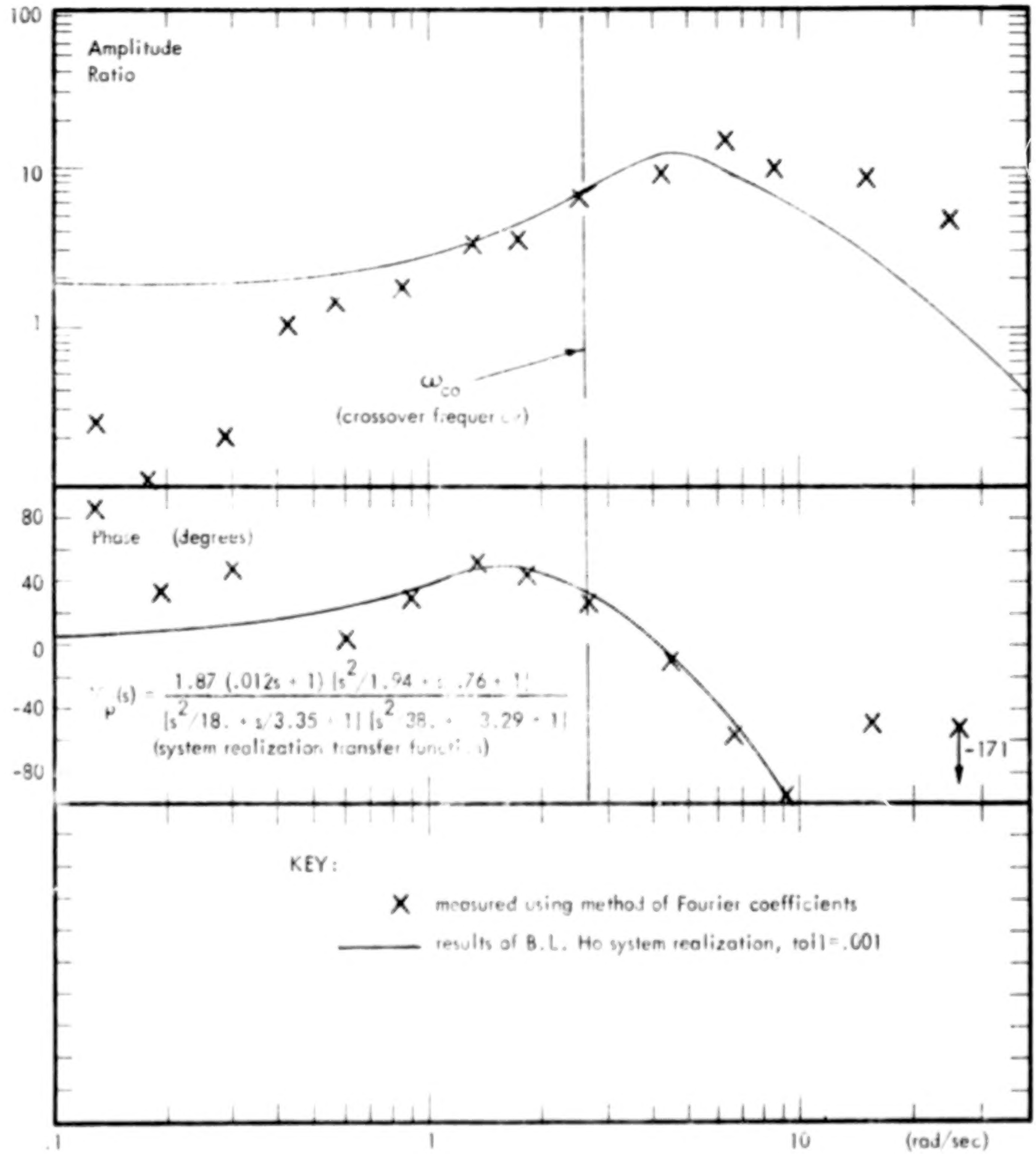


Figure 15.- Frequency response of B.L. Ho system realization for human operator, $Y_c(s) = 1/s^2$

REFERENCES

1. Tustin, A.: "The Nature of the Operator's Response in Manual Control and Its Implications for Controller Design." Journal of the Institute of Electrical Engineers, vol. 94, part IIA, 1947, pp. 190-202.
2. McRuer, D., et al: "Human Pilot Dynamics in Compensatory Systems." AFFDL-TR-65-15, July 1965.
3. McRuer, D., and Krendel, E. S.: "The Human Operator as a Servo System Element." Journal of the Franklin Institute, vol. 267, no. 5, May, and no. 6, June 1959.
4. Obermeyer, R. W., and Muckler, F. A.: "Modern Control System Theory and Human Control Functions." NASA-CR-256, July 1965.
5. Milton, J. L., and Wolfe, F. J.: "Eye Fixations of Aircraft Pilots: Fixations during Zero Reader Approaches in a Jet Aircraft." USAF WADC Tech. Report No. 52-17 (ATI 149050), February 1952.
6. Senders, J. W., et al: "Human Visual Sampling Processes: A Simulation Validation Study." BBN Report No. 1485, January 1967.
7. Carbonell, J. R.: "A Queuing Model of Many-Instrument Visual Sampling;" IEEE Transactions on Human Factors In Electronics, vol. HFE-7, no. 4, December 1966.
8. Baron, S., and Kleinman, D. L.: "The Human as an Optimal Controller and Information Processor." NASA CR-1151, September 1968.
9. Baron, S., et al: "Application of Optimal Control Theory to Prediction of Human Performance in a Complex Task." Proceedings of the 5th NASA-University Conference on Manual Control, March 1969.
10. Baron, S., and Kleinman, D. L.: "The Human as an Optimal Controller and Information Processor." IEEE Transactions on Man-Machine Systems, vol. MMS-10, no. 1, March 1969, pp. 9-17.
11. Kleinman, D. L.: "Optimal Control of Linear Systems with Time Delay and Observation Noise." IEEE Transactions on Automatic Control, vol. AC-14, no. 5, October 1969.
12. Kleinman, D. L.: "Optimal Stationary Control of Linear Systems with Control Dependent Noise." IEEE Transactions on Automatic Control, vol. AC-14, no. 6, December 1969.
13. Kleinman, D. L.: "On the Stability of Linear Stochastic Systems." IEEE Transactions on Automatic Control, vol. AC-14, no. 4, August 1969.

14. Kleinman, D. L. et al: "An Optimal Control Model of Human Response Part I: Theory and Validation." Automatica, vol. 6, no. 3, May 1970.

15. Falb, P. I., and Kovatch, G.: "Dynamical System Modeling of Human Operators: Preliminary Report." 2nd NASA-University Annual Conference on Manual Control, NASA SP-128, February 1966.

16. Ho. Y. C., and Whelan, B. H.: "An Approach to the Identification and Control of Linear Dynamic Systems with Unknown Parameters." IEEE Transactions on Automatic Control, vol. AC-8, July 1963, pp. 255-256.

17. Kalman, R. E.: et al: Topics in Mathematical System Theory, McGraw-Hill Book Co., 1969.

18. Englar, T. S.: "Identification of Linear Systems." To be published as a NASA Contractor Report, 1970.

19. Shirley, R. S.: "Application of a Modified Fast Fourier Transform to Calculate Human Operator Describing Functions." NASA-TM-X-1762, March 1969.

Section 4

ENGINE CONTROL AND ADVANCED PROPULSION TECHNIQUES

NORMAL SHOCK AND RESTART CONTROLS FOR A SUPERSONIC AIRBREATHING PROPULSION SYSTEM

By George H. Neiner, Gary L. Cole, and Francis J. Paulovich
Lewis Research Center

SUMMARY

Results of experimental investigations to control a supersonic mixed-compression inlet coupled to a turbojet engine are presented. Special instrumentation and servoactuators were developed to have sufficiently fast dynamic response that basic propulsion system dynamics were the main limitation to controllability. In some cases servoactuator input signals were electronically limited to simulate moderate performance flight hardware.

An inlet normal shock position control system was developed utilizing a two loop control. It utilized an outer loop controller which provided proportional-plus-integral action with a complex equalization refinement. The performance of this control is demonstrated for the case of the inlet terminated with a close-coupled choked orifice, a choked orifice downstream of a large volume, and a turbojet engine. The controlled inlet was subjected to diffuser exit airflow disturbances. Significant reduction of disturbance induced shock motion over the frequency range of 0 to 40 Hz was demonstrated. Reduction of an inherent inlet resonance at 55 Hz was not accomplished by the control system tested. Little change in closed-loop performance was produced by the various termination impedances.

Two schemes for cross-coupling inlet and engine controls were investigated. In one case control bypass door flow area was used as a primary, fast-acting variable for control of inlet normal shock position, while engine speed was biased through an integral control action to return the doors to a low drag operating position. In the other case engine speed was used as the primary variable for control of the inlet's normal shock position, while the control bypass doors were reset through an integral action to return engine speed to its commanded value. The potential advantage here is to use the inherently fast engine fuel throttle for shock control, thus relieving the high response requirement of bypass door actuation. Both cross-coupled controls are demonstrated to return the propulsion system quickly to design conditions after disturbances in diffuser exit airflow.

A control system was tested for restarting an inlet after an inlet unstart. It featured closed-loop control of normal shock position during unstarted operation. This type of control was demonstrated to minimize pressure distortion at the diffuser exit and maximize inlet pressure recovery during the restart cycle.

INTRODUCTION

To achieve maximum efficiency in a started mixed-compression inlet, it is desirable to maintain the normal shock as near the throat as possible. This maximizes the inlet pressure recovery and minimizes total pressure distortion at the compressor face. Downstream and upstream disturbances can displace the shock and cause decreased pressure recovery and increased distortion, if the normal shock moves downstream. If the disturbance causes the normal shock to move upstream of the inlet aerodynamic throat, the inlet will unstart. To prevent either of these undesirable conditions from happening it is necessary to regulate normal shock position with active control. Under normal conditions the normal shock cannot remain at a fixed position in the converging section of the inlet just upstream of the throat. As a result, once the normal shock moves upstream of the inlet throat, it is expelled from the inlet. Inlet unstart is characterized by either very low pressure recovery when the normal shock is in a stable condition at the inlet entrance, or in an unstable condition known as buzz. Buzz results in nonsteady shock motion and airflow due to oscillations of the shock system ahead of the inlet. This results in very rapid pressure and flow pulsations within the inlet. This, in turn, can seriously affect engine operation or result in structural damage to the airframe. Normal shock position can be changed by changing either overboard bypass airflow or changing engine speed, thereby changing engine airflow. Conventional normal shock control systems use overboard bypass doors, located near the compressor face, operating in a closed loop utilizing aerodynamic signals. In the case of the restart control, manipulation of the contraction ratio (throat area to inlet flow capture area) by means of a collapsing or translating centerbody is also required.

The performance of present inlet normal shock control systems is limited primarily by the dynamics of the bypass doors and aerodynamic sensors. Three of the four control systems presented in this paper utilize high-performance bypass doors and aerodynamic sensors. The purpose of using these high-performance bypass doors was to evaluate controls whose performance was determined primarily by the inherent inlet dynamics, rather than control hardware.

Three normal shock controls and one restart control are presented in this paper. The first normal shock control utilized only inlet manipulated parameters for control of normal shock position as is done in present day design practice. It utilized high response bypass doors, being unconventional in this regard.

The second control presented is an inlet-engine cross-coupled normal shock control. For this control, the previously mentioned high-performance bypass doors were utilized to manipulate bypass airflow. This was the primary variable. A secondary variable engine speed, through a bypass door error signal, was changed to obtain a more efficient propulsion system operating point.

The third control presented is also an inlet-engine cross-coupled normal shock control. In this instance engine speed was used as the primary variable regulating normal shock position. Bypass airflow was used as a secondary variable. In this case the bypass door response was made slower than the engine speed loop response.

The last control presented is one which is intended for emergency use in case of inlet unstart. In this event it is desirable for the propulsion system control to promptly restore inlet and engine to normal operation. The restart control presented utilized a combination of scheduled and closed loop control to provide a rapid return to normal operation.

SYMBOLS

A_8	engine exhaust nozzle area [cm^2]
BPD	bypass doors
CB	centerbody
H	total pressure [N/cm^2]
$K_{c56}G_{c56}$	controller transfer function that uses the throat exit static pressure for feedback [V/V]
$K_{c92}G_{c92}$	controller transfer function that uses the diffuser exit static pressure for feedback [V/V]
$K_{1,2,3,4,5}$	various controller gain factors
M	Mach number
m	mass flow [kg/sec]
N	engine speed [rpm]
P	static pressure [N/cm^2]
R_c	cowl lip radius [23.7 cm]
S	Laplace operator [1/sec]
T	total temperature [K]
VG	compressor variable geometry
W	corrected airflow [kg/sec]
W_o	inlet capture airflow [kg/sec]
W_2	engine airflow [kg/sec]
w_f	engine fuel flow [kg/sec]
X	position [cm]
Δ	indicates incremental change in variable

Subscripts:

bc	control bypass doors
bd	disturbance bypass doors
bias	additional signal which modifies command signal
c	control
cb	centerbody
cl	cowl lip
com	command
de	diffuser exit
eng	engine
max	maximum
min	minimum
s	normal shock
sn	fuel spray nozzle
th	throat
0	freestream
2	compressor face station
56	throat exit static pressure tap located 56.13 cm aft of the cowl lip
92	diffuser exit static pressure tap located 92.20 cm aft of the cowl lip

APPARATUS AND PROCEDURE

Inlet Description

The inlet selected for this investigation was an axisymmetric, translating centerbody type. The design Mach number was 2.5 with 60 percent of the total supersonic area contraction occurring internally. The cowl lip diameter of the inlet was 46.5 cm. The inlet had a capture area of 1,760 cm^2 and a design capture corrected airflow of 16.2 kg/sec.

Figure 1 is an isometric drawing of the inlet. The translating centerbody is hydraulically actuated and electronically controlled. The aft portion of the diffuser is compartmented by three struts which extend aft to the compressor face station. In each of the three compartments are located two overboard bypass doors. These doors are hydraulically actuated and electronically controlled.

Boundary layer bleed holes were located in the vicinity of the throat on both the cowl and centerbody surfaces. The cowl bleed exits are shown in Figure 1. The centerbody bleed flow was ducted through two of the three centerbody struts to the freestream. Vortex generators were used on the centerbody of the inlet. The overboard and ejector bypass exits are also shown in the figure. The ejector bypass flow was used to cool the engine

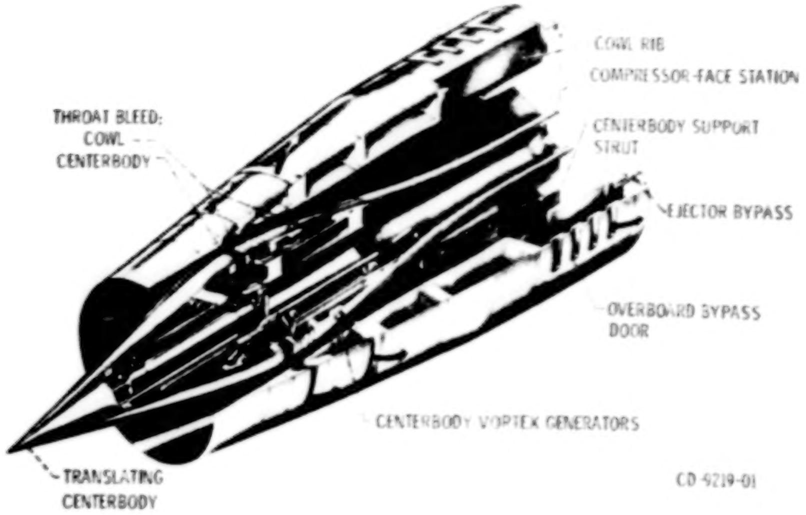
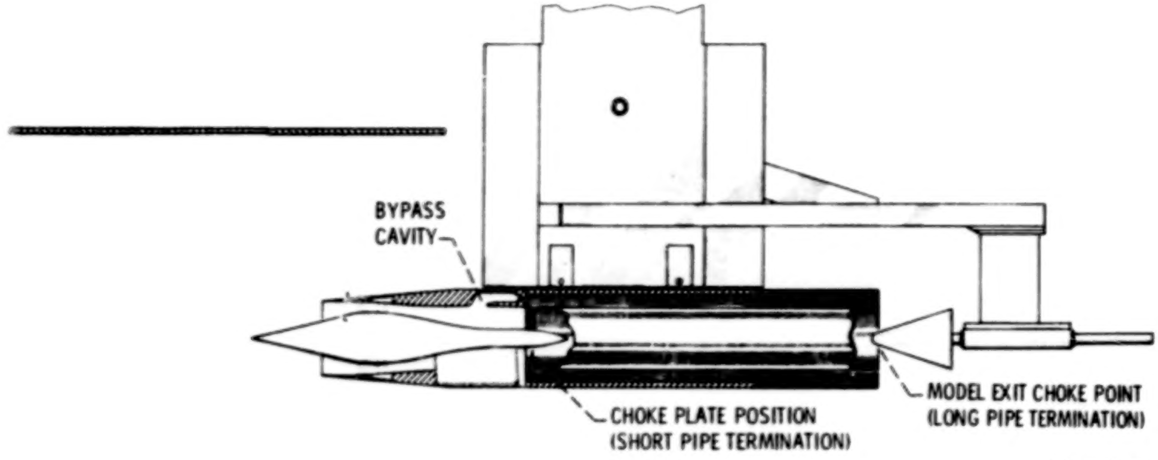

COWL RIB
COMPRESSOR-FACE STATION
THROAT BLEED:
COWL
CENTERBODY
CENTERBODY SUPPORT STRUT
EJECTOR BYPASS
OVERBOARD BYPASS DOOR
CENTERBODY VORTEX GENERATORS
TRANSLATING CENTERBODY
CD-9219-01

Figure 1.- Isometric view of inlet model

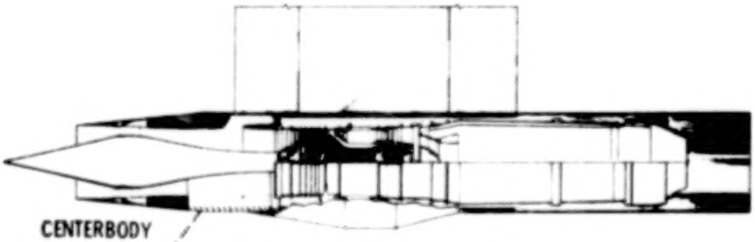
when the inlet was terminated with the engine. At design conditions it passes 3 percent of the inlet mass flow. Additional steady-state performance and dynamics of the inlet can be found in references 1 to 3.

Inlet Terminations

Three inlet terminations were used for this investigation; they are shown schematically in Figure 2. The terminations were installed in a cylindrical nacelle 63.5 cm in diameter. The first termination was a long cold-flow pipe, choked at the model exit plug and shown in Figure 2(a). The cold pipe length was 236 cm, while its diameter was 38.0 cm. The exit plug area was remotely adjustable. The second termination consisted of the choke plate and is also shown in Figure 2(a). It was positioned 146.5 cm downstream of the cowl lip. The flow area of the choke plate was 653 cm^2. When the choke plate was installed, the model exit plug area was increased to cause a choked flow condition at the plate. The third termination was the J85-13 turbojet engine as shown in Figure 2(b), whose inlet guide vanes were 152.2 cm from the cowl lip.


BYPASS
CAVITY
CHOKE PLATE POSITION
(SHORT PIPE TERMINATION)
MODEL EXIT CHOKE POINT
(LONG PIPE TERMINATION)
CD-9739-01

(A) INLET COLD-PIPE CONFIGURATIONS


CENTERBODY
BLEED STRUT
CD-9740-01

(B) INLET-ENGINE CONFIGURATION

Figure 2.- Test configurations

Test Set-Up

Figure 3 is a schematic representation of the inlet controls experiment as installed in the 10- by 10-ft Supersonic Wind Tunnel at Lewis Research Center. A small ±10 V general purpose analog computer located in the control room was used for signal conditioning, programming controllers, and other real time calculations. Close-coupled dynamic pressure transducers were utilized for data and control. Their frequency response was flat to approximately 250 Hz.

Those signals used for control are indicated on Figure 3. The high response normal shock controls used static pressures P_{56} and P_{92} as feedback signals to manipulate the inlet bypass doors to control the normal shock. P_{56} is a static pressure

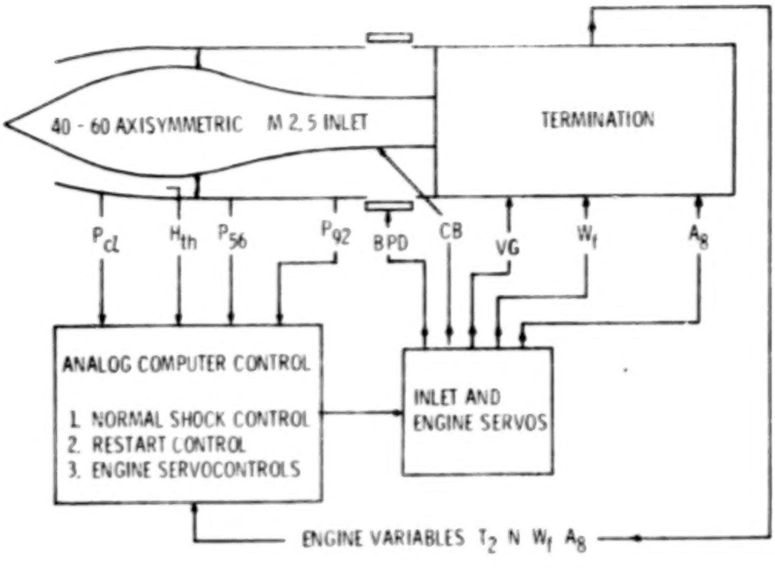

40 - 60 AXISYMMETRIC M 2.5 INLET
TERMINATION
P_{cl}
H_{th}
P_{56}
P_{92}
BPD
CB
VG
W_f
A_8
ANALOG COMPUTER CONTROL
1. NORMAL SHOCK CONTROL
2. RESTART CONTROL
3. ENGINE SERVOCONTROLS
INLET AND ENGINE SERVOS
ENGINE VARIABLES T_2 N W_f A_8

Figure 3.- Schematic representation of the inlet-engine controls experiment in the 10- by 10-foot supersonic wind tunnel

at the throat exit location, this tap being 56.13 cm aft of the cowl lip. P_{92} is a static pressure in the diffuser region, 92.20 cm aft of the cowl lip. Three symmetrically spaced bypass doors were used for control, and the other three bypass doors were used to generate a downstream flow disturbance. Each door was independently controllable by means of an individual electrohydraulic servomechanism. The bypass door servos were designed to operate sinusoidally to 140 Hz, and exhibited flat response to 100 Hz.

The restart control used the ratio of a cowl lip static pressure to a throat total pressure to determine if the inlet was started or unstarted. These are designated in Figure 3 as P_{cl} and H_{th}. When the inlet unstarted, H_{th} decreased due to a reduction in inlet total pressure recovery, and P_{cl} increased due to a transition from supersonic to subsonic flow. In these circumstances the ratio P_{cl}/H_{th} increased significantly and was used as an indicator of inlet unstart. If the inlet was unstarted, a signal was generated to extend the centerbody (CB) to restart the inlet.

When the engine termination was used, the engine inlet temperature (T_2), engine mechanical speed (N), engine fuel flow (w_f), and engine exhaust nozzle area (A_8) variables were fed back to the computer. These signals along with P_{56} were used in various phases of the three programs to manipulate or implement the engine servos: engine variable compressor and bleed geometry (VG), engine fuel flow (w_f) and engine exhaust nozzle area (A_8).

Procedure

Both steady state and dynamic data were taken. The dynamic data consisted of frequency responses and step transients. The criterion for evaluating the high response normal shock control systems was the minimization of disturbance induced shock motion. The disturbance doors were oscillated sinusoidally to 140 Hz to provide a symmetric downstream disturbance.

For the cross-coupled controls various inlet and engine signals were monitored during the transients. The purpose of the first cross-coupled inlet-engine control was to maintain minimum overboard bypass airflow by resetting engine speed while maintaining a desired shock position with the high response inlet bypass doors. The control was subjected to a step disturbance in diffuser exit corrected airflow. The purpose of the second cross-coupled inlet-engine control was to maintain normal shock position control by manipulating engine fuel flow and then resetting the inlet bypass doors to maintain a desired engine operating speed. This control was also subjected to a step change in diffuser exit corrected airflow.

The restart controls were evaluated by monitoring transients of the various inlet signals throughout the restart cycle. The criteria were minimization of inlet distortion and maximization of pressure recovery throughout the restart. A pulse-type disturbance from the three disturbance doors produced the inlet unstart.

RESULTS AND DISCUSSION

Normal Shock Control

Present day inlet control system performance is limited primarily by the dynamics of the bypass doors and aerodynamic sensors. The high response normal shock position control presented in this paper utilized high response bypass doors and aerodynamic sensors. The purpose of this investigation was to evaluate controls whose performance was limited by the inherent inlet dynamics rather than the control hardware.

Normal shock controls were first investigated analytically. Prior to the experimental tests, experimental inlet open loop dynamics were not available. An analog simulation of the wave equations was used to obtain an analytical open loop inlet frequency response. This was then curve fitted and put into transfer function form. This transfer function, along with the control hardware transfer functions, which were obtained from curve fitting experimental data, were then used with the root locus design technique to determine the various controllers. The most promising controllers were then tried experimentally. The results of the best experimental controls are presented next. Figure 4 shows the general controller block diagram. It is a two loop control. The outer loop has P_{56} as its feedback signal. The inner loop uses ΔP_{92} as its feedback signal. ΔP_{92} is generated by feeding back the P_{92} signal through a high pass filter.

Figure 5 shows the responses of throat exit static pressure P_{56} to a disturbance in bypass door corrected airflow W_{bd}. The squares and circles show the open-loop response of the inlet when terminated with the cold pipe and engine, respectively. It was determined (ref. 3) that the open-loop responses of the inlet with choke plate and engine terminations were quite similar. The closed-loop responses of the inlet when terminated by the cold pipe, choke plate, and engine are shown as the solid, dashed, and broken curves, respectively. These curves have been normalized to their open-loop values at 1 Hz. The controller gains which correspond to $K_{c56}G_{c56}$ and $K_{c92}G_{c92}$ from Figure 4 are indicated in the legend of Figure 5. $K_{c56}G_{c56}$ and $K_{c92}G_{c92}$ are the controller transfer functions that use the throat exit and diffuser exit static pressures respectively for feedback. The engine termination closed-loop response was run at a slightly lower outer loop gain which accounts for its slightly poorer response. The closed-loop responses show that at 1 Hz the disturbance induced shock motion has been reduced by over 90 percent compared to the open-loop response. Also, the disturbance induced shock motion has been reduced below that of the open loop out to 40 Hz. The results of the closed-loop responses also indicate that the performance of the two-loop, compensated integral control is relatively unaffected by the inlet termination. This is desirable, since the input impedance of turbojet engines can vary as a function of their operating point. Additional information on the high response normal shock controls investigation is given in reference 4.

Cross-Coupled Normal Shock Controls

Case I, bypass airflow as primary variable, engine speed as secondary variable. - Diffuser exit corrected airflow is equal to the sum of the bypass (disturbance and control) and engine corrected airflows. Engine airflow is a function of engine rotor speed. Thus, by changing either engine speed or bypass door position the inlet's normal shock position can be changed. Conventionally, the inlet shock position loop and the engine speed

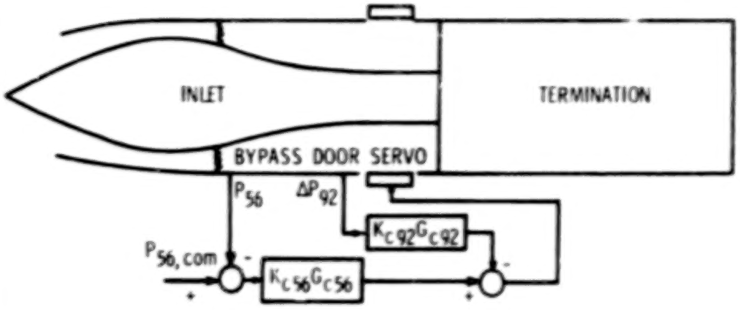

INLET
TERMINATION
BYPASS DOOR SERVO
P_{56}
ΔP_{92}
$P_{56,com}$
$K_{c92}G_{c92}$
$K_{c56}G_{c56}$
+
-
+
-

Figure 4.- Normal shock controller block diagram-compensated integral controller

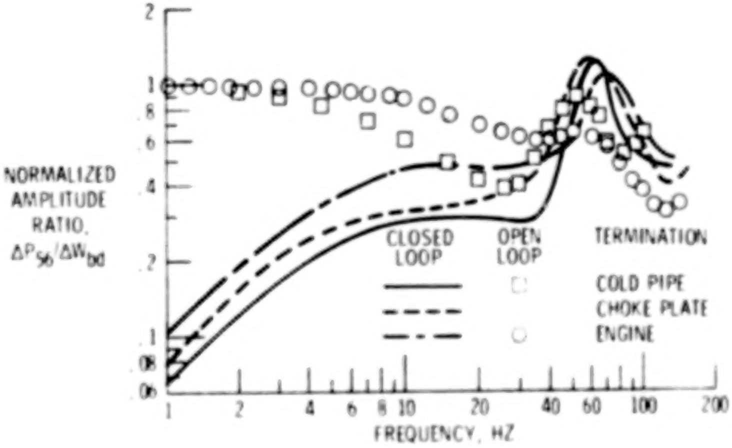
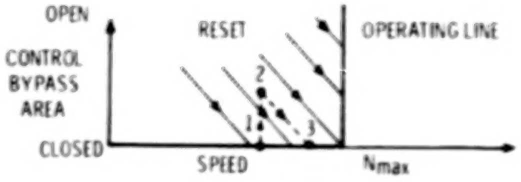
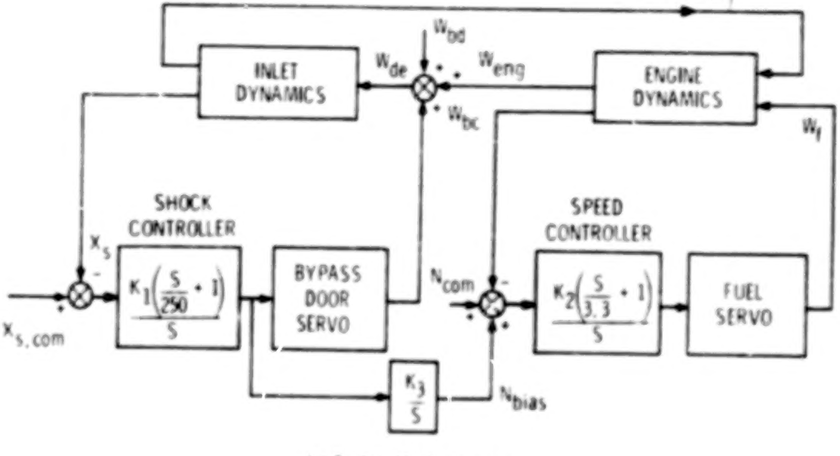


Figure 5.- Experimental response of throat exit static pressure to downstream airflow disturbance for a multiloop normal shock position control using three inlet terminations


OPEN
CONTROL BYPASS AREA
CLOSED
RESET
OPERATING LINE
1
2
3
SPEED
N_{max}

(b) Control action.


P_t
W_{bd}
INLET DYNAMICS
W_{de}
W_{eng}
ENGINE DYNAMICS
W_{bc}
W_f
SHOCK CONTROLLER
X_s
$X_{s,com}$
$\frac{K_1\left(\frac{S}{250}+1\right)}{S}$
BYPASS DOOR SERVO
SPEED CONTROLLER
N_{com}
$\frac{K_2\left(\frac{S}{3.3}+1\right)}{S}$
FUEL SERVO
$\frac{K_3}{S}$
N_{bias}

(a) System block diagram.

Figure 6.- Cross-coupled inlet-engine control: Case I - bypass airflow primary variable. Engine speed secondary variable for control of normal shock position

loop function independently. One possibility for an improved cross-coupled control is to leave the conventional loops intact, but feed a speed biasing signal to the engine speed control. This would slowly change the engine speed command signal to force the bypass doors to a position of low spillage, thereby reducing drag. This technique could have a positive effect on propulsion system operating efficiency.

Figure 6(a) shows the schematic of how the controls were interconnected with the inlet and engine dynamics. A previously developed high response normal shock position control is shown as the loop on the left-hand side of the figure. A shock position error signal (actually an error in P_{56}) acting through a proportional plus integral controller was used to manipulate the control bypass doors to control normal shock position. The speed loop is shown in the lower right-hand corner of Figure 6. Here a speed error signal acts through a proportional-plus-integral controller to manipulate fuel flow to control speed. As shown in the figure, both the bypass doors and speed acting through appropriate inlet-engine dynamics affects normal shock position. Propulsion system efficiency was improved by adding a slow integral reset action to engine speed which allowed the overboard bypass spillage to be reduced.

Figure 6(b) shows the control action. The ordinate is control bypass door area and the abscissa is engine speed. The chosen operating line is shown as the heavy line. Thus, under quiescent conditions either the bypass doors will be closed or engine speed will be at its maximum value. When a disturbance occurs the normal shock position control manipulates the control bypass doors to maintain fixed shock position. As the reset action changes engine speed, the bypass doors move as necessary to maintain constant shock position. It is noted that increasing engine speed will require decreasing control bypass door area for constant shock position. Thus, diagonal lines on the diagram will correspond to values of speed and control bypass door area for constant position of the normal shock.

If a diffuser exit airflow disturbance causes the normal shock to move forward in the throat, the normal shock is brought back to its desired position when the high response normal shock control opens the control bypass doors. This is illustrated in Figure 6(b) by going from point 1 to point 2. Then the slower reset action adds an N_{bias} to the speed command signal telling the engine to speed up. As the engine speeds up, the control bypass doors close to keep the normal shock at its desired position. This path is illustrated in Figure 6(b) by going from point 2 to point 3. For any size disturbance the system returns to the operating line.

Figure 7 shows the transient response of several system variables to a step change in disturbance bypass door area for this control. Starting from the top of the figure, the variables are disturbance bypass door position feedback voltage, control bypass door position feedback voltage, throat exit static pressure P_{56}, fuel spray nozzle pressure, and engine mechanical speed. The normal shock position control quickly compensates for the disturbance and the normal shock does not move very far as evidenced by the small change in P_{56}. The traces show that engine speed is slowly increased, thus allowing the control bypass doors to be brought back to their initial closed position.

Case II, engine speed as primary variable, bypass airflow as secondary variable. - After having tested the previous cross-coupled control, an entirely different approach was taken in controlling normal shock position. In this instance normal shock position is controlled through fuel flow manipulating engine speed. While this normal shock position control is slower than the ones previously discussed, it demonstrates the capability of a normal shock position control using a much less sophisticated bypass door system. A slow reset action was added to the above mentioned normal shock position control which would change the bypass door flow, thus allowing the engine speed to return to its initial commanded value.

Figure 8(a) shows a schematic of how the controls were interconnected with the inlet and engine dynamics. This control also used the same speed loop that was discussed in the previous cross-coupled control. As shown in the left-hand side of Figure 8(a), a shock position error signal (actually P_{56} error) acts through a proportional-plus-integral controller to generate an N_{bias} signal which changes the speed command signal to effect normal shock position control. A slower integral reset action resetting the control bypass doors was added to the normal shock position control. This reset the bypass doors until engine speed again equaled N_{com} or until a bypass door position limit was reached.

Figure 8(b) shows the control action. The ordinate is control bypass door area and the abscissa is engine speed. The chosen operating line is shown as the heavy line. While maintaining normal shock position control, the reset action takes place in the direction of the diagonal arrows until the system reaches the operating line. For this control the N_{com} line chosen must be less than N_{max} to allow for normal shock position control for forward disturbance induced shock motion.

If a diffuser exit airflow disturbance causes the normal shock to move upstream, the engine speed is increased through the action of the shock controller, thus bringing the normal

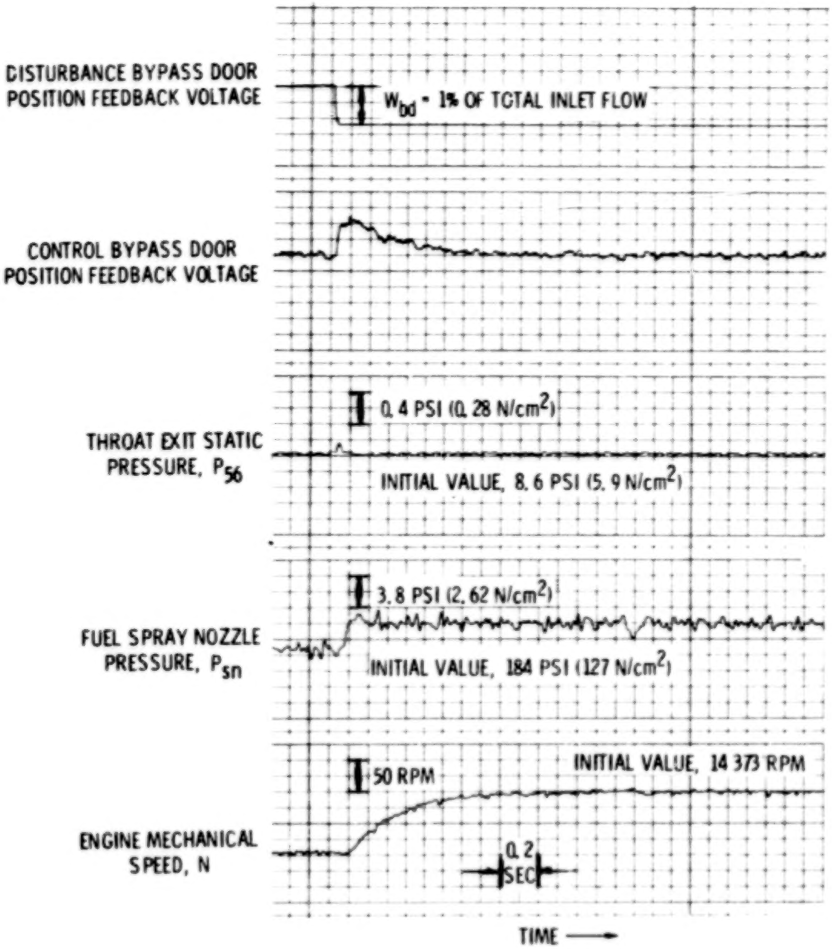
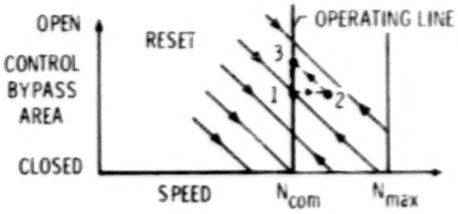
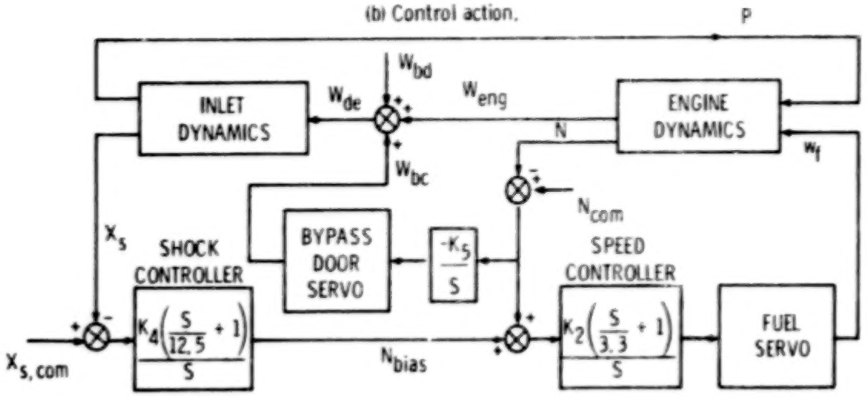


Figure 7.- Transient response of cross-coupled inlet-engine control: Case I - bypass airflow primary variable, engine speed secondary variable for control of normal shock; step decrease disturbance in diffuser exit airflow


OPEN
CONTROL BYPASS AREA
CLOSED
RESET
OPERATING LINE
3
1
2
SPEED
N_{com}
N_{max}

(b) Control action.


P
W_{bd}
W_{de}
W_{eng}
INLET DYNAMICS
ENGINE DYNAMICS
N
w_f
W_{bc}
N_{com}
X_s
SHOCK CONTROLLER
BYPASS DOOR SERVO
$\frac{-K_5}{S}$
SPEED CONTROLLER
$\frac{K_4\left(\frac{S}{12.5}+1\right)}{S}$
$\frac{K_2\left(\frac{S}{3.3}+1\right)}{S}$
FUEL SERVO
$X_{s,com}$
N_{bias}

(a) System block diagram.

Figure 8.- Cross-coupled inlet-engine control: Case II - engine speed primary variable, bypass airflow secondary variable for control of normal shock position

shock back to its initial desired position. This action is illustrated in Figure 8(b) by going from point 1 to point 2. The reset action of the control bypass doors opens the bypass doors thus allowing engine speed to be decreased in order to maintain the normal shock at its desired position. This reset action takes place until engine speed reaches N_{com} or until a bypass door limit is reached. This action is illustrated by going from point 2 to point 3 in Figure 8(b).

Figure 9 shows the transient response of several system variables to a step change in disturbance bypass door area for the control just discussed. Starting from the top of the figure, the variables are disturbance bypass door feedback voltage, control bypass door feedback voltage, fuel valve position feedback voltage (which is proportional to fuel flow), throat exit static pressure P_{56} and engine mechanical speed. Normal shock motion, as expected, is more than that for the previously discussed cross-coupled control. After the disturbance, fuel valve position is increased, thus increasing engine speed to bring P_{56} back to its initial value. At this point the (N_{com}-N) error signal slowly opens the control bypass doors, thus allowing

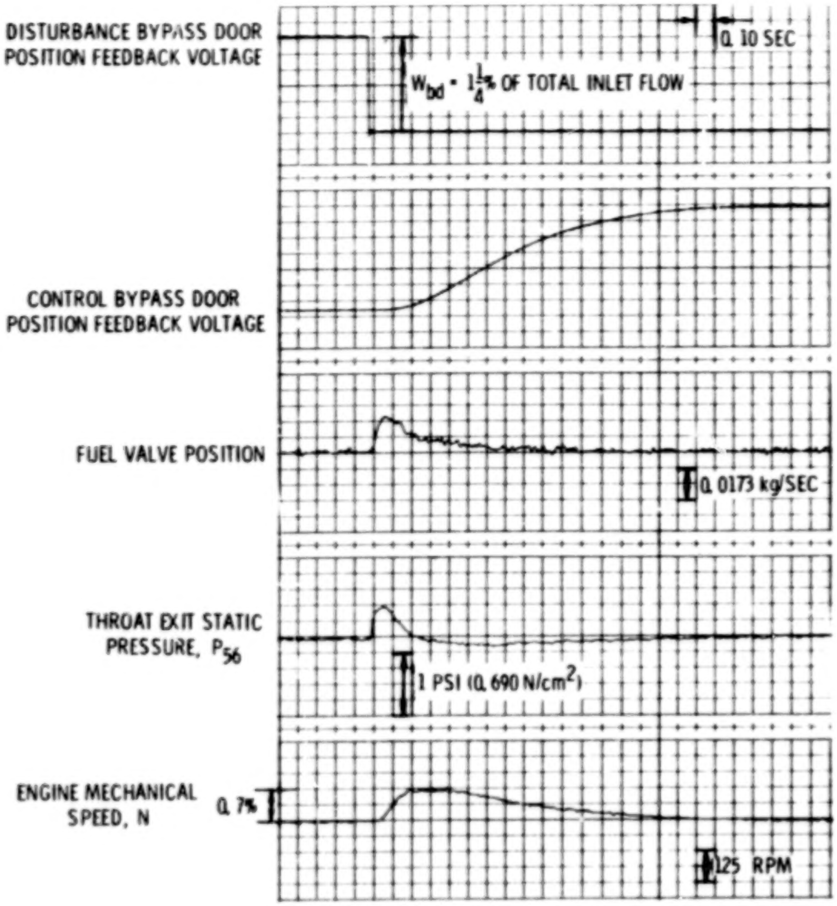


Figure 9.- Transient response of cross-coupled inlet-engine control: Case II - engine speed primary variable, bypass airflow secondary variable for control of normal shock position

engine speed to return to its commanded value, while the normal shock is kept at its desired position.

Restart Control

Although the control systems previously discussed provide effective regulation of the started inlet, additional capabilities are required in the event of an inlet unstart. The phenomenon of inlet unstart is accompanied by a drastic decrease in inlet pressure recovery. Under normal circumstances a normal shock position control closes the bypass doors in response to a decrease in pressure recovery. If the unaltered shock position control executes this type of action after an inlet unstart, it would contribute to inlet buzz, an unstable unstarted condition characterized by oscillatory flow.

Conventional restart controls disable the normal shock control system and open the bypass doors upon sensing an inlet unstart. The technique stabilizes the inlet by choking the throat. Under these conditions a secondary shock forms downstream of the throat. Concurrently, the centerbody is translated or collapsed to increase the ratio of throat area to cowl lip flow area (contraction ratio) which allows reswallowing of the normal shock. The normal shock control is reactivated sometime after the inlet is restarted.

To insure choking of the throat, the bypass door opening should be of sufficient magnitude to allow for the reduced engine airflow which would result from an engine stall. Although this would allow successful restarting of the inlet, excessive bypass door opening would cause high distortion levels in the inlet, particularly subsequent to restart when the normal shock would reach an extremely supercritical location. In addition, the pressure recovery during the restart cycle would be less than that potentially attainable.

Figure 10 shows the inlet total pressure recovery and distortion versus engine plus overboard bypass mass-flow ratio for various centerbody positions. Approximately 60 steady-state total pressure probes were located just in front of the engine face. H_{max} was the maximum individual total pressure. H_{min} was the minimum individual total pressure. H_2 was the average of the 60 total pressures. The tailed symbols are the peak recovery points. With the exception of diamond symbols on the distortion plot, the data show that distortion increases and pressure recovery decreases as the normal shock is moved to more supercritical locations for both started and stable unstarted inlet operation. If the engine is operating and the inlet is either started or stable unstarted, it is possible to stall the engine if the distortion is high enough. Thus, it would be desirable to operate with the normal shock (whether it is the primary or secondary normal shock) as close to the throat as possible without causing buzz or repeated inlet unstart.

An investigation was conducted to develop a restart control which retained closed-loop control of the bypass doors throughout the restart sequence.

Figure 11 shows a schematic representation of the restart control which was investigated. The normal shock control system utilized the compensated integral controller, previously discussed. During started operation of the inlet, the command to the normal shock controller was scheduled as a function of centerbody position.

The ratio of the cowl lip static to throat total was used to detect unstart. If this signal exceeded a predetermined value, the comparator relay switched to the unstart state. The command to the shock control was then determined by the unstarted schedule. This schedule commanded a value of throat exit static pressure which was sufficiently low to be attainable with the inlet unstarted. At the same instant, the centerbody was extended at its maximum slewing rate to increase the contraction ratio. During the translation of the centerbody, the throat exit static pressure command was scheduled to increase in a stepwise continuous manner.

The instant restart was sensed, the comparator returned the shock control command to the started schedule and reversed the travel of the centerbody. For each centerbody position the started command schedule was also increased in a stepwise continuous manner. This prevented excessively supercritical shock positions during the cycle. Automatic positioning of the bypass doors during the entire restart cycle was accomplished with this system.

A typical unstart transient, followed by a controlled restart cycle, is shown in Figure 12. The arrows indicate increasing values of the variables. For this test the inlet was terminated with the large volume with a choked exit.

The normal shock was initially positioned on the verge of unstart by the closed-loop bypass control. A pulse-type decrease in disturbance door area initiated the inlet unstart. The initial increase in control door area indicates the attempt of the normal shock control to prevent the unstart. Unable to prevent the unstart, and responding to the decreasing pressure, the normal shock control began to close the control doors, which is undesirable.

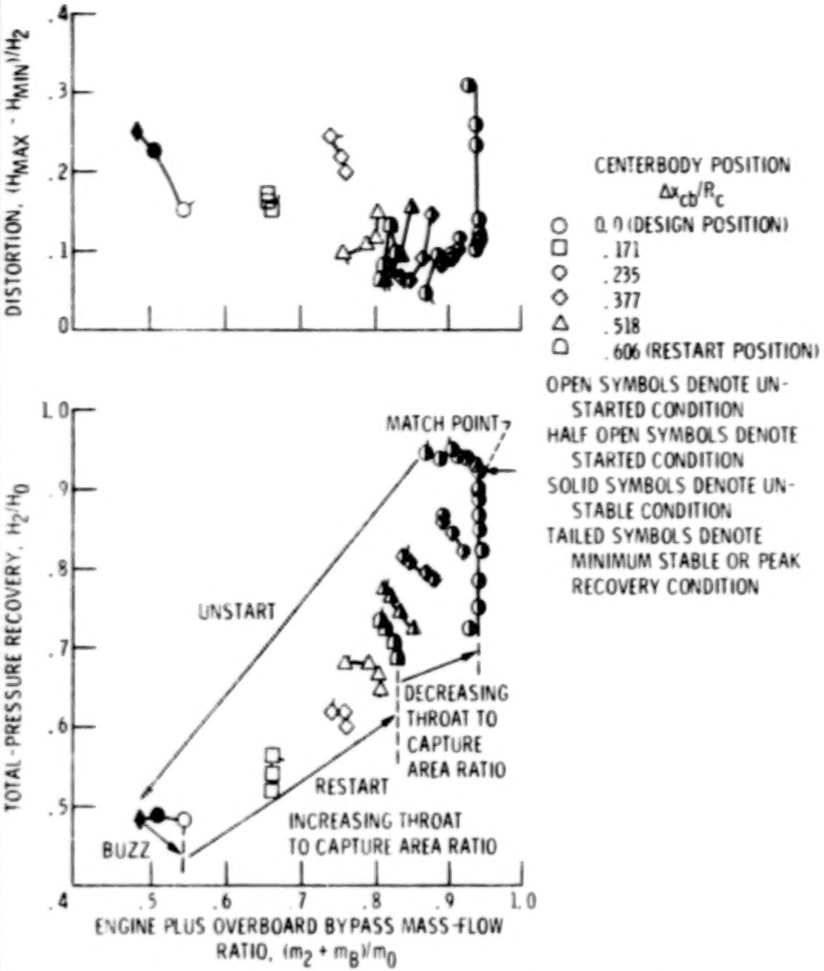


Figure 10.- Inlet total pressure recovery and distortion for various centerbody positions. Free-stream Mach number, 2.50; free-stream temperature, 317°K; Reynolds number, 3.82×10^6; ejector area to inlet capture area ratio, 0.0081; choked exit corrected airflow, 15.83 kilograms per second. Data from reference 5

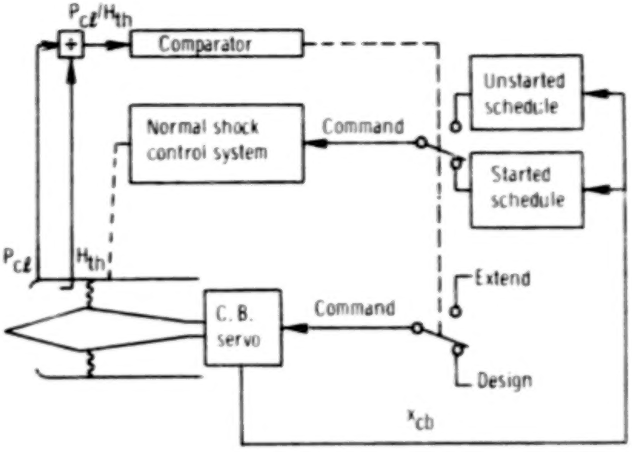

$P_{c\ell}/H_{th}$
+
Comparator
Unstarted schedule
Normal shock control system
Command
Started schedule
$P_{c\ell}$
H_{th}
Extend
C. B. servo
Command
Design
x_{cb}

Figure 11.- Schematic of restart control

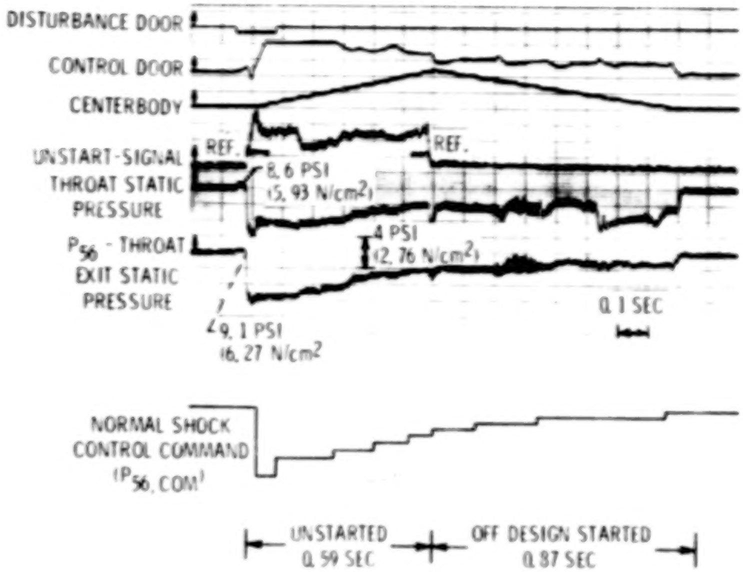


Figure 12.- Controlled restart transient

When the unstart signal exceeded the reference value, the normal shock control command signal was switched to the unstarted schedule. This caused the control door area to increase to attain the lower commanded value of throat exit static pressure.

At the same time the centerbody was extended, the control door area remained full open until the measured value of throat exit static pressure, P_{56}, dropped below the commanded value. The normal shock controller then adjusted control door area to attain the scheduled command pressure. The centerbody was extended until restart was indicated by the unstart signal dropping below the reference value. At that time the centerbody travel was reversed.

Upon restart, the normal shock first formed between the throat static and P_{56} taps. This was indicated by the large instantaneous drop in the throat static pressure and a relatively small drop in P_{56}. The normal shock controller then followed the started schedule of the command signal. The control door area was decreased, moving the terminal shock upstream of the throat static pressure. As the centerbody returned to design, the normal shock again moved downstream of the throat static pressure, because the inlet throat was fixed relative to the centerbody. With the centerbody back on design, the normal shock returned to the operating point as indicated by the pressure levels in the inlet. The total restart cycle took 1.46 sec and was limited by the centerbody slewing rate.

The same control was also tested with the inlet terminated with the J85-13 turbojet. Inlet unstarts with the engine were conducted at Mach numbers of 2.5 and 2.0. In these tests the inlet unstart occurred in less time than that in which a servo could react. The consequences to the engine depended on the magnitude of the pressure decrease and on the initial operating value of steady state compressor pressure ratio. The results found in these tests are summarized in Table I.

In these tests the inlet was satisfactorily restarted in all cases independent of the condition of the engine. For this initial investigation, if the combustor flamed out, no attempt was made to return the engine to normal operation either during or after restarting the inlet.

In a second investigation, attempts were made to recover normal engine operation after inlet restart. In this case a conventional hydromechanically controlled engine was used. Since no manipulation of the power lever was exercised, the engine controller held a constant ratio of fuel flow to compressor inlet static pressure. To allow for the case of combustor blowout the spark ignition was turned on just before the inlet unstart and left on until the engine relit. It was observed that the compressor stall always cleared immediately after the unstart transient. In all cases the combustor relit. The relight time took about 15 sec from the time of unstart if the inlet was just stabilized but left unstarted. For this condition, after the inlet was unstarted, the centerbody was left at its design position and the bypass doors were opened to a larger area to stabilize the inlet shock system. After the unstart transient, the inlet pressure recovery was 0.75. The relight time took only about 5 or 6 sec from the time of unstart if the inlet was restarted and brought back to its design operating point conditions. In this instance the inlet operating point pressure recovery was 0.88. It took 0.5 sec to restart the inlet and bring it back to its operating point conditions. It is felt that more ambitious manipulation of engine parameters could produce improvements in combustor relight times over these preliminary results.

CONCLUSIONS

The performance characteristics of normal shock control systems which utilized high performance bypass doors were investigated. Significant reduction of disturbance induced shock

TABLE I.- SUMMARY OF INLET-ENGINE UNSTART TESTS

Tunnel Mach number, M_0	Inlet pressure recovery,[a] percent	Initial compressor pressure	Design compressor pressure ratio	Engine corrected speed, percent	Resulting engine operation
2.5	94 → 24	4.0	4.0	85.0	Stall and flameout- no recovery
2.0	94 → 70	4.9	4.4	86.3	Stall and flameout- no recovery
2.0	94 → 70	4.5	4.4	86.3	Stall and no flameout- engine recovered
2.0	94 → 70	4.2	4.4	86.3	No stall or flameout

[a] Pressure recovery before inlet unstart - pressure recovery immediately after inlet unstart.

motion was achieved over the frequency range 0 to 40 Hz with a two-loop, compensated integral controller. Analyses and design of this system required consideration of the inherent inlet dynamics to 100 Hz. Reduction of the inlet resonance at 55 Hz was not accomplished by the control systems investigated. Adequate control of the normal shock was maintained, whether the inlet was terminated by a large volume cold pipe, choke plate at the engine face, or a J-85 turbojet engine.

The feasibility of a high response normal shock position control coupled to a slower resetting action that reset engine speed was demonstrated. By resetting the engine speed, it was possible to reduce overboard bypass door airflow which, in turn, reduced drag due to overboard bypass airflow. Experimental results show the control to operate satisfactorily, returning the inlet quickly to design conditions when subjected to step disturbances in inlet diffuser exit corrected airflow.

Controlling normal shock position through engine speed was demonstrated. A bypass door reset action was added to this normal shock position control loop. The purpose of this reset action was to make it possible to maintain a commanded engine speed, while also maintaining a commanded shock position. The entire control was tested against a step disturbance in inlet diffuser exit airflow. The control quickly brought the normal shock and engine speed back to their commanded values.

The ability of the normal shock control system to control the secondary shock in the unstarted inlet led to the development of a restart control which featured closed-loop positioning of the bypass doors. This was achieved by scheduling the normal shock control command as a function of centerbody position for both started and unstarted conditions. This type of control minimized distortion and maximized pressure recovery during the restart cycle.

Restart cycle times on the order of 1.5 sec were possible with this control, this time being determined primarily by the slewing velocity of the centerbody servo.

It is felt that significant further work in inlet controls could employ integrated inlet-engine-flight control with the use of on-line digital computer controllers. Simultaneous manipulation of inlet and engine variables could be employed to further improve propulsion system operating efficiency and immunity to disturbances.

REFERENCES

1. Cubbison, Robert W.; Meleason, Edward T.; and Johnson, David F.: "Performance Characteristics from Mach 2.58 to 1.98 of an Axisymmetric Mixed-Compression Inlet System with 60-Percent Internal Contraction." NASA TM X-1739, 1969.

2. Cubbison, Robert W.; Meleason, Edward T.; and Johnson, David F.: "Effect of Porous Bleed in a High-Performance Axisymmetric, Mixed-Compression Inlet at Mach 2.50." NASA TM X-1692, 1968.

3. Wasserbauer, Joseph F.: "Dynamic Response of a Mach 2.5 Axisymmetric Inlet with Engine or Cold Pipe and Utilizing 60 Percent Supersonic Internal Area Contraction." NASA TN D-5338, 1969.

4. Crosby, Michael J.; Neiner, George H.; and Cole, Gary L.: "Restart and High Response Terminal Shock Control for an Axisymmetric Mixed-Compression Inlet with 60 Percent Internal Contraction." NASA TM X-1792, 1969.

5. Cole, Gary L.; Neiner, George H.; and Crosby, Michael J.: "An Automatic Restart Control System for an Axisymmetric Mixed-Compression Inlet." NASA TN to be published.

BLANK PAGE

OPTIMAL CONTROL OF A SUPERSONIC INLET TO MINIMIZE FREQUENCY OF INLET UNSTART

By Bruce Lehtinen, John R. Zeller, and Lucille C. Geyser
Lewis Research Center

SUMMARY

This paper describes a preliminary investigation into the use of modern control theory for the design of controls for a supersonic inlet. In particular, the task of controlling a mixed-compression supersonic inlet is formulated as a linear optimal stochastic control and estimation problem. An inlet can exhibit an undesirable instability (unstart) due to excessive inlet normal shock motion. For the optimal control formulation of the inlet problem, a non-quadratic performance index, which is equal to the expected frequency of inlet unstarts, is used. This physically meaningful performance index is minimized for a range of inlet disturbance and measurement noise covariances.

A lumped parameter, single input-single output inlet model is considered, where the disturbance is a white Gaussian airflow perturbation at the compressor face station. The measured variable is normal shock position, corrupted by additive Gaussian white measurement noise. A Kalman filter is used to generate the estimates of the state variables. The linear controller gains are found by solving a series of quadratic control problems (Riccati equation) using the quadratic equivalence principle. The variation of the performance index with noise and quadratic performance index weighting parameters is investigated. Mean square behavior of the control signal is presented to examine possible control saturation problems.

INTRODUCTION

For optimum performance of present-day supersonic aircraft, it is necessary to provide an efficient means for decelerating the air from supersonic to subsonic velocity before it enters the compressor of the turbojet engine. The most common device for accomplishing this is a variable geometry supersonic inlet as shown in Figure 1. The type shown is an axisymmetric mixed compression inlet to be used on future supersonic transports. In normal operation, air enters the inlet past a weak oblique shock wave and is compressed supersonically past the minimum area point (throat) until the terminal normal shock. After the normal shock, the air is further compressed subsonically until it reaches the inlet of the turbojet engine (compressor face). The job of the inlet is to maximize the recovered pressure at the compressor face over a wide range of operating conditions and to maintain a stable flow pattern despite external disturbances.

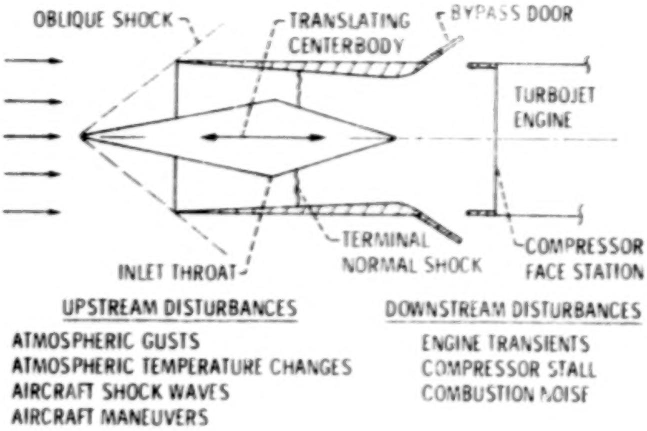

OBLIQUE SHOCK
TRANSLATING CENTERBODY
BYPASS DOOR
TURBOJET ENGINE
TERMINAL NORMAL SHOCK
COMPRESSOR FACE STATION
INLET THROAT
UPSTREAM DISTURBANCES
ATMOSPHERIC GUSTS
ATMOSPHERIC TEMPERATURE CHANGES
AIRCRAFT SHOCK WAVES
AIRCRAFT MANEUVERS
DOWNSTREAM DISTURBANCES
ENGINE TRANSIENTS
COMPRESSOR STALL
COMBUSTION NOISE

Figure 1. - Schematic of supersonic axisymmetric inlet.

The flow configuration of this type of inlet is bistable; the desired started condition just described; and a second unstarted condition having a strong external shock system, subsonic flow in the initial convergent section, poor pressure recovery, and sometimes oscillatory flow (buzz). The desired started condition becomes unstable and switches to unstart when either the Mach number at throat approaches 1 or when the terminal normal shock moves upstream of the minimum flow area (throat). Such an unstart occurring in flight will not only interact with the engine, producing compressor stall and combustor flameout, but the increased nacelle drag will cause a sudden yawing of the aircraft. In addition, a started condition with the terminal normal shock too far downstream of the throat will give relatively poor pressure recovery and produce undesirable inlet flow distortions at the engine face.

Control is, therefore, required to maintain throat Mach number and shock position within acceptable limits, while maintaining efficient inlet operation. Basic control parameters are bypass doors and a translating inlet centerbody. Opening the bypass doors allows air to be dumped overboard, causing the shock to move downstream away from the throat. The translating centerbody varies the throat area, thereby varying the throat Mach number. A combination of these two modes is used to ensure stable operation in the face of upstream and downstream disturbances.

The inlet can encounter both random and deterministic disturbances in a typical flight environment. Deterministic-type disturbances are those such as atmospheric pressure changes due to a shock wave from a passing aircraft, pilot-induced engine transients, aircraft maneuvers, and engine compressor stall. Disturbances which are random in nature are those such as atmospheric gusts, atmospheric temperature changes, and combustion noise fed back from the engine.

Present designs (refs. 1, 2) are based on minimizing system response to deterministic disturbances such as ramp or sinusoidal changes in upstream Mach number and/or engine air flow rate. Another approach is discussed by Barry (ref.3). Using a statistical model for atmospheric turbulence, he developed a method for predicting the expected number of unstarts per flight mile and compared inlet controls on the basis of their ability to minimize unstart frequency.

The approach described in this paper is to formulate the inlet control problem as a linear stochastic optimal control problem using, as the performance index, the expected frequency of unstarts. This design method is direct in that it produces a control which minimizes a physically meaningful performance index rather than some intermediate quantity such as closed-loop frequency response or mean-square shock position. The nature of this performance index dictates unique characteristics for the optimum controller otherwise not required.

SYMBOLS

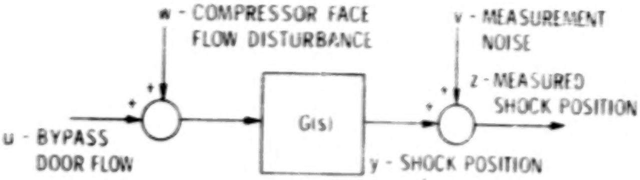
A	system matrix
B	control matrix
D	plant disturbance matrix
G(s)	inlet transfer function
H	measurement matrix
N	quadratic weighting matrix on state-control
P_c	quadratic weighting matrix on control
p_c'	control weighting parameter
Q	quadratic weighting matrix on state
Q_c	spectral density of plant disturbance
R	quadratic weighting matrix on state
s	Laplace variable
t	time
u	control vector

v	measurement noise vector
W	free parameter
w	plant disturbance vector
x	state vector
y	output vector (shock position)
z	measurement vector (measured shock position)
α	shock position tolerance
$\delta(\tau)$	delta function
λ	expected frequency of unstarts
λ_{eq}	equivalent quadratic performance index
σ_y^2	mean square shock position
$\sigma_{\dot{y}}^2$	mean square shock velocity
τ	time
$\delta()$	first variation of
$E()$	expected value of
Superscripts:	
$\cdot$	differentiation with respect to time
T	tranpose operation

OPTIMAL CONTROL FORMULATION

To avoid unstarts due to inlet downstream disturbances requires manipulation of bypass doors. Fast-acting doors have been used in our experimental programs in this regard. To avoid unstarts due to inlet upstream disturbances requires varying the throat area. For this function, a slow translating centerbody was used. In addition, the response of unstart to upstream disturbances is very fast. The combination of very fast system response and slow control capability makes active control of throat area to avoid unstarts due to upstream disturbances either very difficult to mechanize or very limited in effectiveness. For these reasons, our initial studies have concentrated on the optimum control of bypass doors in the presence of inlet downstream disturbances.

The inlet control problem considered herein is the single input-single output case shown in Figure 2. The dynamic performance of a representative supersonic inlet is modeled by the transfer function shown in Figure 2. The disturbance to the inlet w is modeled as white Gaussian noise with a spectral density Q; w represents a random air flow perturbation at the compressor face. The output y (shock position) is measured through a noisy channel with white Gaussian measurement noise v having a spectral density R.



INLET DYNAMICS

$$G(s) = \frac{y(s)}{u(s)} = \frac{e^{-0.01s}}{\left(\frac{s}{40} + 1\right)\left[\left(\frac{s}{420}\right)^2 + \frac{2(0.707)}{420}s + 1\right]}$$

PLANT DISTURBANCE COVARIANCE

$$E(w(t)\,w(t+\tau)) = Q\delta(\tau)$$

MEASUREMENT NOISE COVARIANCE

$$E(v(t)\,v(t+\tau)) = R\delta(\tau)$$

Figure 2. - Single-input single-output inlet control.

The performance index which is to be minimized is given by

$$\lambda = \frac{1}{2\pi}\sqrt{\frac{\sigma_{\dot{y}}^2}{\sigma_y^2}}\ \exp -\left(\frac{\alpha^2}{2\sigma_y^2}\right) \qquad (1)$$

where

λ = expected frequency of unstarts

σ_y^2 = mean square value of shock position

$\sigma_{\dot{y}}^2$ = mean square value of shock velocity

α = distance between undisturbed shock position and the inlet throat.

Equation (1) is a classical exceedance equation (ref. 4) that gives the expected frequency with which the Gaussian random variable y will exceed the level α. The control problem is to find a control on u, given the measurement z, which will minimize λ.

There are two major departures from the usual formulation of the optimum control problem: (1) While both σ_y^2 and $\sigma_{\dot{y}}^2$ are quadratic in the state variables, the performance index λ is not, Eq. (1); (2) The plant to be controlled is basically of distributed parameter form ($e^{-0.01s}$ on Figure 2).

Linear Optimal Control Solution

To make use of the well-known solution to the linear stochastic control and estimation problem for quadratic performance indices (for example, ref. 5), a quadratic approximation to the non-quadratic performance index of Eq. (1) was obtained. This follows the technique employed by Skelton et al (ref. 6) which they term quadratic equivalence.

The performance index of Eq. (1) can be expressed in quadratic form by first taking the variation of λ and setting it equal to zero, obtaining:

$$\delta\left(\sigma_{\dot{y}}^2\right) + \frac{\sigma_{\dot{y}}^2}{\sigma_y^2}\left(\frac{\alpha^2}{\sigma_y^2} - 1\right)\delta\left(\sigma_y^2\right) = 0 \tag{2}$$

Defining

$$W = \frac{\sigma_{\dot{y}}^2}{\sigma_y^2}\left(\frac{\alpha^2}{\sigma_y^2} - 1\right) \tag{3}$$

the following quadratic performance index can be obtained:

$$\lambda_{eq} = E\left[\frac{1}{2}\int_0^\infty \left(x^T Q_c x + 2x^T N u + u^T P_c u\right) dt\right] \tag{4}$$

where

$$\dot{x} = Ax + Bu + Dw \tag{5}$$

and

$$y = Hx \tag{6}$$

$$z = Hx + V \tag{7}$$

is the state variable representation of the single input-single output inlet model of Figure 2 and the quadratic weights are given by:

$$\left.\begin{aligned} Q_c &= A^T H^T HA + WH^T H \\ N &= A^T H^T HB \\ P_c &= B^T H^T HB + p_c' \end{aligned}\right\} \qquad (8)$$

An arbitrary control weighting p_c' has been added.

In terms of the system input u and shock position output y, the quadratic problem to be solved is to make

$$E(Wy^2 + \dot{y}^2 + p_c' u^2) = \text{a minimum}$$

Variable parameters are the control weighting p_c', and the free parameter W. By quadratic equivalence, the non-quadratic problem can be solved by solving the quadratic problem, for W equal to a constant, described by Eqs. (4) through (8) with the stipulation that at the optimum, Eq. (3) must be satisfied. The approach taken in this paper is to simply solve the quadratic problem for various values of W and conduct a search for the optimum by calculating λ each time. The result is then a set of estimator and linear controller gains which minimize the non-quadratic performance index λ.

The distributed parameter term in the plant model was described a third-order Padé approximation as follows:

$$e^{-0.01s} \cong \frac{\left(\frac{-s}{464} + 1\right)\left[\left(\frac{s}{508}\right)^2 - \frac{2(0.724)}{508}\, s + 1\right]}{\left(\frac{s}{464} + 1\right)\left[\left(\frac{s}{508}\right)^2 + \frac{2(0.724)}{508}\, s + 1\right]} \qquad (9)$$

The state variable formulation of the plant was derived, replacing $e^{-0.01\,s}$ on Figure 2 with Eq. (9). The control problem in having a plant time delay corresponds to one having a plant with zeroes in the right-half plane

Control Computations

The method used to solve the inlet control problem and the computations required are outlined in Figure 3. Program inputs

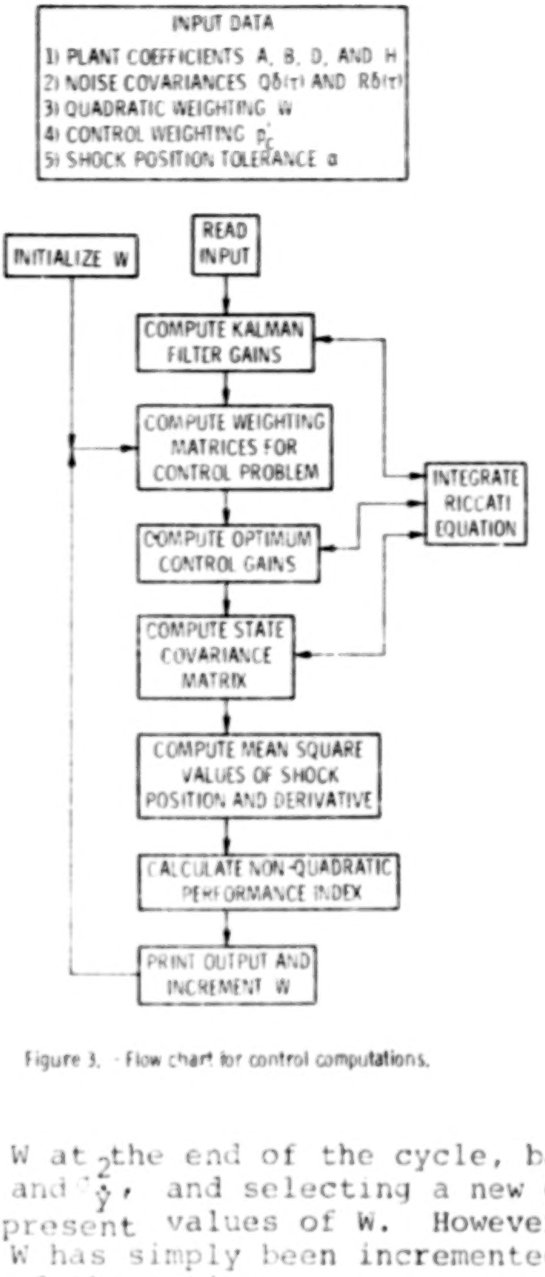

INPUT DATA
1) PLANT COEFFICIENTS A, B, D, AND H
2) NOISE COVARIANCES Qδ(τ) AND Rδ(τ)
3) QUADRATIC WEIGHTING W
4) CONTROL WEIGHTING p'c
5) SHOCK POSITION TOLERANCE α
INITIALIZE W
READ INPUT
COMPUTE KALMAN FILTER GAINS
COMPUTE WEIGHTING MATRICES FOR CONTROL PROBLEM
INTEGRATE RICCATI EQUATION
COMPUTE OPTIMUM CONTROL GAINS
COMPUTE STATE COVARIANCE MATRIX
COMPUTE MEAN SQUARE VALUES OF SHOCK POSITION AND DERIVATIVE
CALCULATE NON-QUADRATIC PERFORMANCE INDEX
PRINT OUTPUT AND INCREMENT W

Figure 3. - Flow chart for control computations.

are plant coefficient matrices A, B, D, and H, noise spectral densities Q and R, quadratic weight. W, control weight p_c', and controlled variable tolerance α.

The first step in the solution is to compute the Kalman filter gains, which requires the steady-state solution of the Riccati equation. The steady-state solution is obtained by a straightforward numerical integration. Then, for a given value of W, the control quadratic weighting matrices Q_c, N, and P_c are calculated. Using these weights, the control problem is solved for the optimum feedback gain matrix, where again, a Riccati equation is solved.

Using the computed filter and feedback gains, the steady-state solution for the state covariance matrix is obtained, which is then used to obtain the mean square values of the shock position and its derivative (σ_y^2 and $\sigma_{\dot{y}}^2$). Finally, the performance index λ is calculated, W incremented, and the cycle repeated. The search procedure could be automated by computing W at the end of the cycle, based on the computed values of σ_y^2 and $\sigma_{\dot{y}}^2$, and selecting a new guess for W based on past and present values of W. However, for the results obtained thus far, W has simply been incremented in a step-wise fashion in search of the optimum.

RESULTS

Shock Position Control to Minimize Unstarts

A state variable model for the inlet was obtained most conveniently in phase variable form. A sixth-order model was used, being comprised of the third-order denominator of G(s) (Figure 2) and the third-order Padé approximation for the plant dead time given in Eq. (9).

The solution procedure outlined in Figure 3 was followed for various combinations of control weighting, plant and measurement noise covariances and shock position tolerances to determine the relative influences these parameters have on the expected frequency of unstarts. Figure 4 shows some of the results obtained for a unity control weighting ($p_c' = 1$) and a plant to measurement noise ratio of 50 ($Q/R = 50$). λ, unstarts per second, is plotted as a function of $\sqrt{W}$ for values of α^2/Q which are considered typical for the inlet being used.

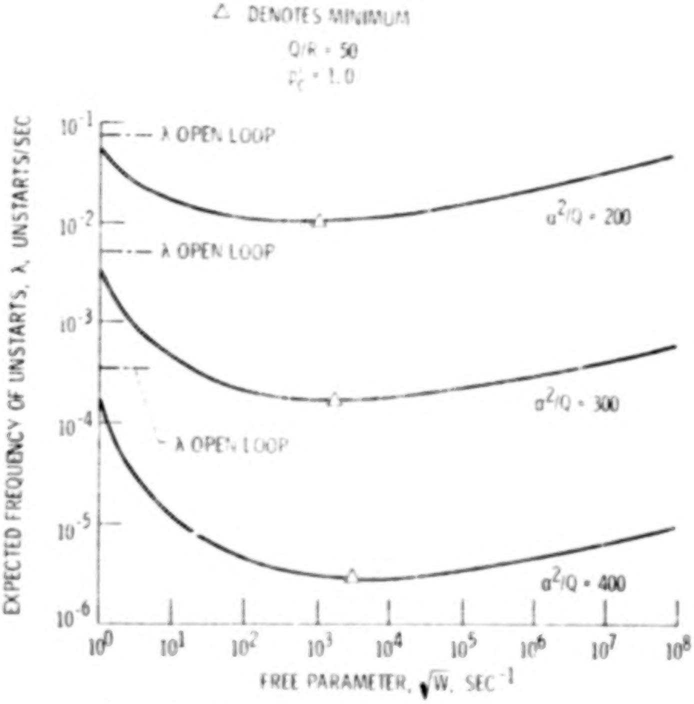


Figure 4. - Expected frequency of unstarts versus free parameter.

There is a set of controller gains for each value of W which minimizes the expected frequency of unstarts. This occurs at $\sqrt{W} = 10^3$ to 10^4 rad/sec. This frequency is where the control begins cutting off to avoid large values of $\sigma_{\dot{y}}^2$. At values of $\sqrt{W}$ beyond the minimum point, the increase in $\sigma_{\dot{y}}^2$ in Eq. (1) causes λ to increase even though σ_y^2 continues to decrease. This is a special optimum control characteristic resulting from using the performance index λ. Also, the effectiveness of control in preventing unstarts can be seen by comparing the open-loop values of λ for the three curves shown with the corresponding minima. Because of the exponential term in the performance index, it is very sensitive to control. For $\alpha^2/Q = 300$, the open loop gives $\lambda = 5 \times 10^{-3}$/sec, whereas optimum closed loop gives $\lambda = 1.7 \times 10^{-4}$/sec. That is, mean time between unstarts, λ^{-1}, increased from 3.3 minutes to 1.63 hours. Figure 4 also shows the sensitivity of λ to shock position tolerance α. As α is made larger, the unstart frequency decreases. However, large values of α mean that the shock is positioned closer to the engine, causing an overall loss in efficiency due to decreased inlet pressure recovery.

It can also be noted that at a minima, W satisfied Eq. (3). This fact could have been used to iteratively change W in the calculation procedure, recomputing controller gains until a system was obtained for which Eq. (3) was satisfied. As indicated before, this procedure was not followed, but rather, λ vs $\sqrt{W}$ plots as on Figure 4 were made to find the minimum and to investigate the behavior of the performance index.

Additional solutions were obtained for different levels of measurement noise. Figure 5 shows cases where measurement noise level is relatively low (Q/R = 1000). For comparison, the dotted curves show the cases where Q/R is 50 for the same values of the parameter α^2/Q. It is apparent that control does much better with low measurement noise. For $\alpha^2/Q = 300$ the mean time between unstarts is 13.9 hours for the low measurement noise case as compared to 1.63 hours for high measurement noise. The optimum $\sqrt{W}$ increased from 2000 to the very high 20,000 rad/sec. This emphasizes the importance of finding improved shock position sensing methods and low-noise transducers. Figure 5 also shows that a system designed to be optimal for one value of measurement noise spectral density will not remain optimal if measurement noise intensity should change. Fortunately, over the range of conditions on Figure 5, one fixed value of $\sqrt{W}$ (about 2000 rad/sec) could be used. Some performance loss would occur, but only where λ is extremely low.

It is important to know the amount of control effort expended in obtaining optimal control. It may turn out that for certain noise levels, the control effort for optimal control is unreasonable. Figure 6 is a plot of mean square control signal normalized with respect to plant noise. Two curves are shown, one for low and one for high measurement noise. It can be seen that control effort (mean square u, σ_u^2) increases with increasing W and also with increasing measurement noise. The optimum control that minimizes λ for a given Q and R may give a value for σ_u^2 that exceeds the physical limits on the control variable. This condition can be identified by using the curves of Figure 6 to obtain the σ_u^2 required, knowing at what value of W the optimum λ occurs. If σ_u^2 is too high, the control weighting p_c' would be increased and the problem resolved. In this case, the minimum λ would be higher than before, but the larger control weighting would serve to keep the control effort to a reasonable level.

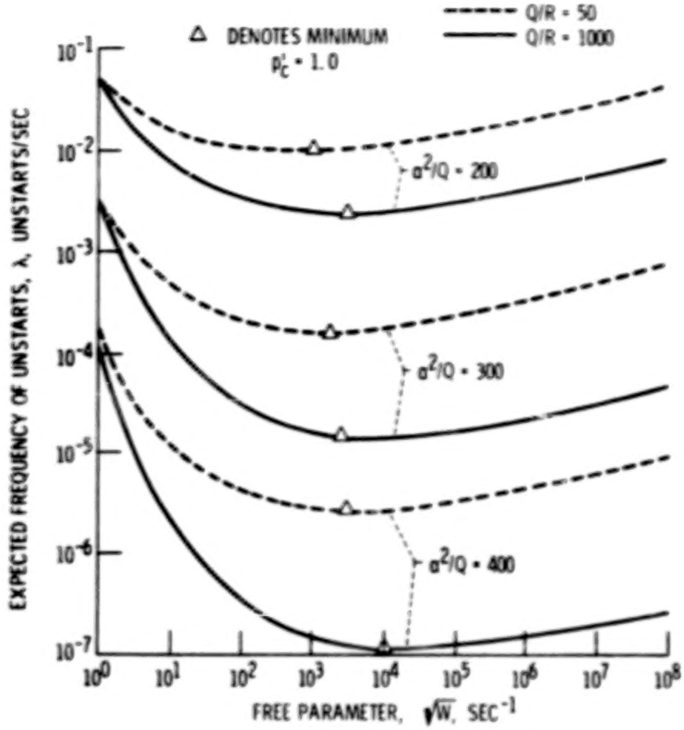


Figure 5. - Expected frequency of unstarts versus free parameter for different measurement noise.

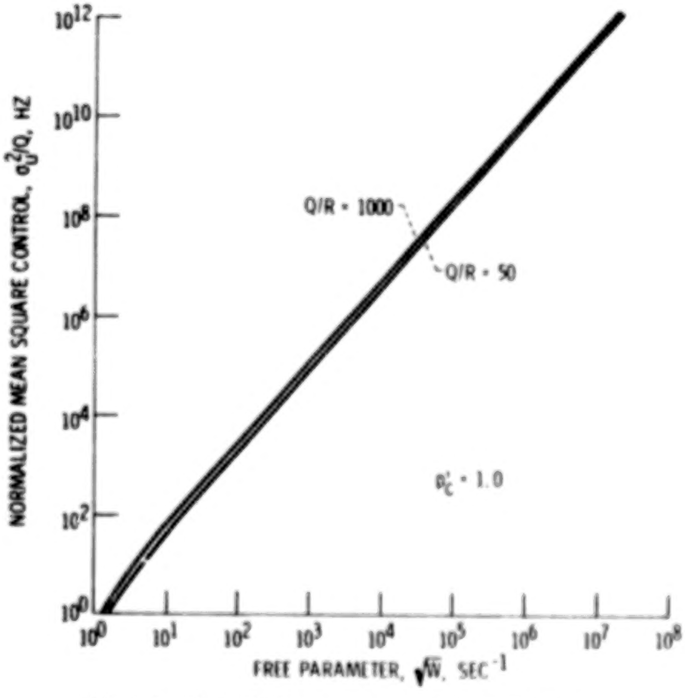


Figure 6. - Normalized mean square control versus free parameter.

Computing Procedures

In solving the control, estimate, and mean square Riccati equations it was found that for the system in phase variable form, numerical problems were encountered if the problem were not time-scaled. After time-scaling, matrix coefficients were more similar in magnitude, greatly improving accuracy and speed of computation.

Actually, to solve for the mean square values (state covariance matrix), the same integration subroutine that was used to solve the Riccati equations was used to integrate the linear-state covariance matrix differential equation to steady state. It was found that this method was quite slow due to the slow convergence of some of the off-diagonal elements. Presently, the mean square equation is being solved as a linear matrix equation, by first partitioning the matrices and then forming a set of $(n^2 + n)/2$ linear equations. For the inlet problem case, $n = 6$, and the linear equation method proved to be two orders of magnitude faster than the integration method.

CONCLUSIONS

This paper has described formulating the problem of controlling a supersonic inlet as an optimal control problem. A physically meaningful non-quadratic performance index, the expected frequency of unstarts, was used in the formulation. A single-input single-output inlet control case was solved via quadratic equivalence by solving a series of quadratic estimation-control problems. The method provides a straightforward procedure for determining a set of linear feedback controls which minimize the non-quadratic performance index. The expected frequency of inlet unstart, using this method for control design, can be minimized. The effects of plant and measurement noise and shock position tolerance were shown to have a large influence on the performance index. However, the free parameter W could be fixed at about 2000 rad/sec and still give almost optimum performance over a wide range of conditions.

The theory used in this paper is valid for multivariable systems. Thus, control of both bypass doors and throat area, with downstream and upstream disturbances, can be included. Provisions can easily be made to allow coloring of both plant and measurement noise. The performance index can be expanded to allow for more than one mode of unstart.

REFERENCES

1. Crosby, M. J., Neiner, G. H., and Cole, G. L.: "High Performance Bypass Control for Mixed-Compression Inlets." Fourth AIAA Joint Propulsion Specialists Conf., June 10-14, 1968.

2. Chun, K. S., and Burr, R. H.: "A Control System for an Axisymmetric Supersonic Inlet." Journal of Aircraft, vol. 6, No. 4, July-August 1969, pp. 306-311.

3. Barry, F. W.: "Development of Atmospheric Gust Criteria for Supersonic Inlet Design." Contract NAS2-4515, Report No. HSER-5195, Hamilton Standard Division, United Aircraft Corp., December 1968.

4. Rice, S. O.: "Mathematical Analysis of Random Noise." In Selected Papers on Noise and Stochastic Processes, ed. by N. Wax, Dover, 1954, p. 61.

5. Bryson, A. E., Ho, Y. C.: Applied Optimal Control, Blaisdell Pub. Co., 1969, pp. 414-417.

6. Edinger, L. D., Skelton, G. B., Stone, C. R., Ward, M.D.; and White, G. D.: "Design of a Load Relief Control System." NASA CR-61169, Honeywell, Inc. April 1967.

BLANK PAGE

CONTROL SYSTEM STUDIES FOR THERMIONIC REACTORS*

By Richard J. Hermsen** and Henrik G. Gronroos***

SUMMARY

In-core thermionic reactor concepts are of interest for space missions that require electrical power in the range of a few tens of kilowatts up to several megawatts. The physical principle involved--thermionic direct conversion of heat to electricity at net efficiencies up to 15 percent--offers potential advantages when compared to other nuclear powerplant concepts. However, the integration of the thermionic diode electrode structure with high-temperature nuclear fuel materials presents new design problems and new reactor physical constraints. Among the topics that must be investigated are those associated with the control system. This paper discusses the results of analytical and simulation studies of thermionic reactor control performed at the Jet Propulsion Laboratory.

INTRODUCTION

Objectives

The control system studies are motivated by the novel character of the thermionic reactor concept, and therefore a relative lack of knowledge of the dynamic behavior. Preceding the control investigations, extensive analytical and analog computer simulation studies of thermionic reactor powerplant open-loop dynamics and stability were undertaken (refs. 1,2). Both high-order, nonlinear and low-order, linear analytical models were developed. The results show that the inherent open-loop reactor response is not consistent with maintaining satisfactory operation over a range of electric loads. Subsequent controller design studies have relied on these data, and have concentrated on investigating the feasibility of constant output voltage control at the reactor terminals. The objectives of these studies may be summarized as follows:

---

* This work presents the results of one phase of research carried out in the Propulsion Research and Advanced Concepts Section of the Jet Propulsion Laboratory, California Institute of Technology, under Contract NAS7-100, sponsored by the National Aeronautics and Space Administration.

** Associate Professor, Electrical and Electronics Department, California State Polytechnic College, Pomona, California.

***Senior Scientist, Jet Propulsion Laboratory, California Institute of Technology.

1. Define the control problem;
2. Investigate the feasibility of constant output voltage control;
3. Investigate applicability of available analytical methods for control system design;
4. Investigate the characteristics of developed controllers.

The analytical studies performed to date represent the initial phase, and have employed linear system equations. However, the linearly derived controller designs have been applied with good results to a non-linear powerplant model. This was done by analog computer simulation, and recently also in an experimental assembly containing liquid metal cooled thermionic diodes (ref.3). In the latter simulation studies only, the neutron kinetics and the controller are programmed on the computer; the rest of the system consists of actual physical elements.

Results

The results from applying two different design approaches are given in this paper. As indicated above, the objective was to obtain a controller for constant output voltage at the reactor terminals. The associated controller for the cesium reservoir loop has also been investigated. This controller maintains the cesium reservoir temperature near optimum and prevents the possibility of thermal runaway, i.e., thermionic burnout (refs 4,5).

In the first analytical approach the state variable feedback technique is used (ref. 6). System states are fed back to obtain a desired linear system transfer function. It is found that reactivity rate, reactivity, neutron density, and output voltage are the important feedback states. To constrain the reactivity rate to design specifications (determined by design and safety considerations), it is necessary to constrain the controller input signal. This constraint slows down the system response considerably. The linear system remains a stable system, which is shown by application of the Popov Criterion. If the reactivity rate itself is constrained, the system becomes unstable for sufficiently large perturbations. With no control signal constraint, loss of reactivity or reactivity-rate feedback causes instability.

The above shortcomings of the results obtained with the state variable feedback method, and the fact that state constraints cannot readily be introduced into the analysis, prompted application of the analytical design method (ref. 7). A performance index minimizing the integral of the sum of the squared error and weighted square of the reactivity rate is selected. The analytical design method allows introduction of constraints on the reactivity rate and also, in an approximate manner, allows consideration of response time, peak overshoot, and steady-state error. Only the voltage is used for controller feedback.

Comparison with the previous state variable feedback results indicates that for the same perturbation the analytical design method gives smaller reactivity rates. To eliminate large reactivity and reactivity rates and yet maintain an acceptable response, a response time of 8 seconds is chosen for the final controller design. Analog computer simulation with linear and non-linear system models were performed to verify the analytical results.

Conclusions

The results of the work performed to date indicate the following general conclusions.

1. Constant output voltage control is feasible. However, depending on the system application some other specific programmed response could also be employed (ref. 8).
2. Existing controller hardware technology is sufficient to meet foreseeable controller performance specifications.
3. A step change in electric load will cause a step change in the voltage at the reactor terminals, and therefore, also at the input to any power conditioner. To avoid this perturbation at the output of the power conditioner, it must be designed to maintain a constant voltage output for the duration of such a transient.

The next step in control system analysis should consider the minimum time response problem with reactivity rate and reactivity constraints. Also the effects of such factors as dead time and dead band should be incorporated. A further analytical development would be to consider the multiple-input/multiple-output problem. A thermionic reactor powerplant has four independent inputs, namely, electric load, cesium reservoir heater power, coolant flow rate, and the reference voltage (or reactivity). The outputs are voltage and current, cesium pressure, coolant temperature, and neutron density. The previous analyses have considered all inputs as disturbances, except the reference signal, and one of the design objectives has been to minimize the effect of such disturbances.

In the following a brief description of the reference thermionic reactor powerplant is given. This is followed by a discussion of the considerations that enter the control problem. Next the results of applying the state variable feedback technique are given. Following this the results obtained with the analytical design method are described.

References 9 and 10 give recent general reviews of thermionic reactor powerplant state-of-the-art and future prospects.

REFERENCE POWERPLANT

Powerplant

For the control system studies in this paper, a reference thermionic reactor powerplant is selected as illustrated by Figure 1. This is a low-power (50 to 100 kWe) fast reactor with fully enriched UC-fuel, and with five series-stacked (flashlight) diodes making up one converter assembly. The coolant is Li-7 in both loops, and the reactor is assumed to be controlled by drums in the reflector. The dc output high-current/low-voltage electrical power at the reactor terminals is fed to a power conditioner for transformation to high dc voltage. The objective is to investigate possible control modes for the reactor and to determine the required characteristics of the controller.

For the purpose of analytical modeling, the reference plant shown in Figure 1 is reduced as illustrated schematically in Figure 2. One of the diodes in a converter assembly is modeled by lumping the parameters. It is assumed that all diodes behave in the same manner in the core. The neutron kinetics is described by the point kinetics equations in the prompt jump approximation (ref. 6). The external coolant loops are also described by lumped parameter equations, and the electric load is purely resistive. A block diagram representation of the powerplant is shown in Figure 3.

In 19th-order, non-linear system of equations is selected as the basic high-order model and simulated on an analog computer. This analytical model has been used for open-loop transient and stability studies, the results of which are reported in references 1 and 2. In general, for the anticipated temperature coefficients of reactivity, the system is open-loop stable and has an inherently slow response.

Analytical Model

The non-linear model is too complicated for effective use in an initial control system study. A reduced fourth order set of linear equations is derived based on the results of the analog computer simulated high-order, non-linear, open-loop system (ref. 11). With the external reactivity as the input and after substitution of parametric values, the reduced set of equations are:

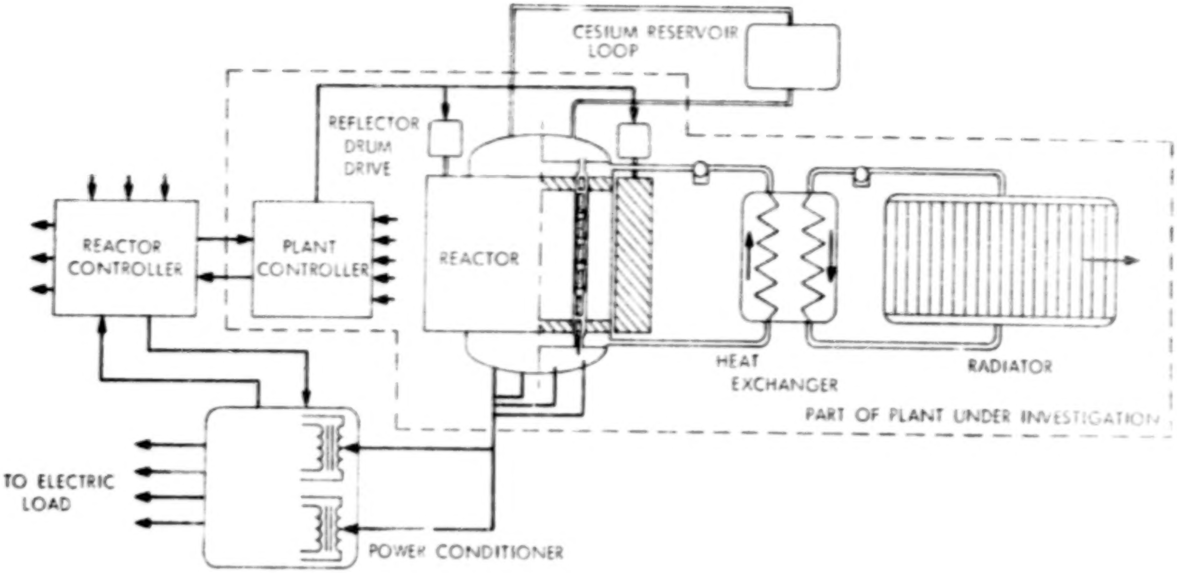

CESIUM RESERVOIR LOOP
REFLECTOR DRUM DRIVE
REACTOR CONTROLLER
PLANT CONTROLLER
REACTOR
HEAT EXCHANGER
RADIATOR
PART OF PLANT UNDER INVESTIGATION
TO ELECTRIC LOAD
POWER CONDITIONER

Figure 1.- Schematic of reference thermionic reactor powerplant

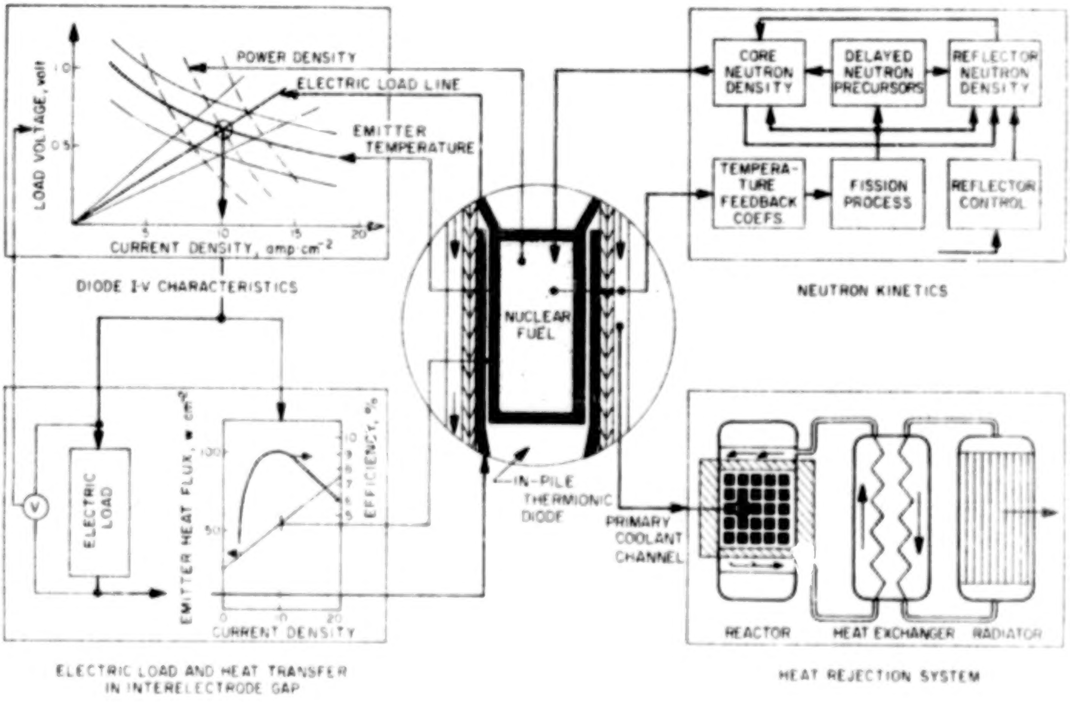

POWER DENSITY
ELECTRIC LOAD LINE
EMITTER TEMPERATURE
LOAD VOLTAGE, volt
1.0
0.5
5
10
15
20
CURRENT DENSITY, amp cm⁻²
DIODE I-V CHARACTERISTICS
CORE NEUTRON DENSITY
DELAYED NEUTRON PRECURSORS
REFLECTOR NEUTRON DENSITY
TEMPERATURE FEEDBACK COEFS
FISSION PROCESS
REFLECTOR CONTROL
NEUTRON KINETICS
NUCLEAR FUEL
IN-PILE THERMIONIC DIODE
PRIMARY COOLANT CHANNEL
V
ELECTRIC LOAD
EMITTER HEAT FLUX, w cm⁻²
EFFICIENCY, %
100
50
10
9
8
7
6
5
0
10
20
CURRENT DENSITY
ELECTRIC LOAD AND HEAT TRANSFER IN INTERELECTRODE GAP
REACTOR
HEAT EXCHANGER
RADIATOR
HEAT REJECTION SYSTEM

Figure 2.- Schematic of simulated thermionic reactor powerplant

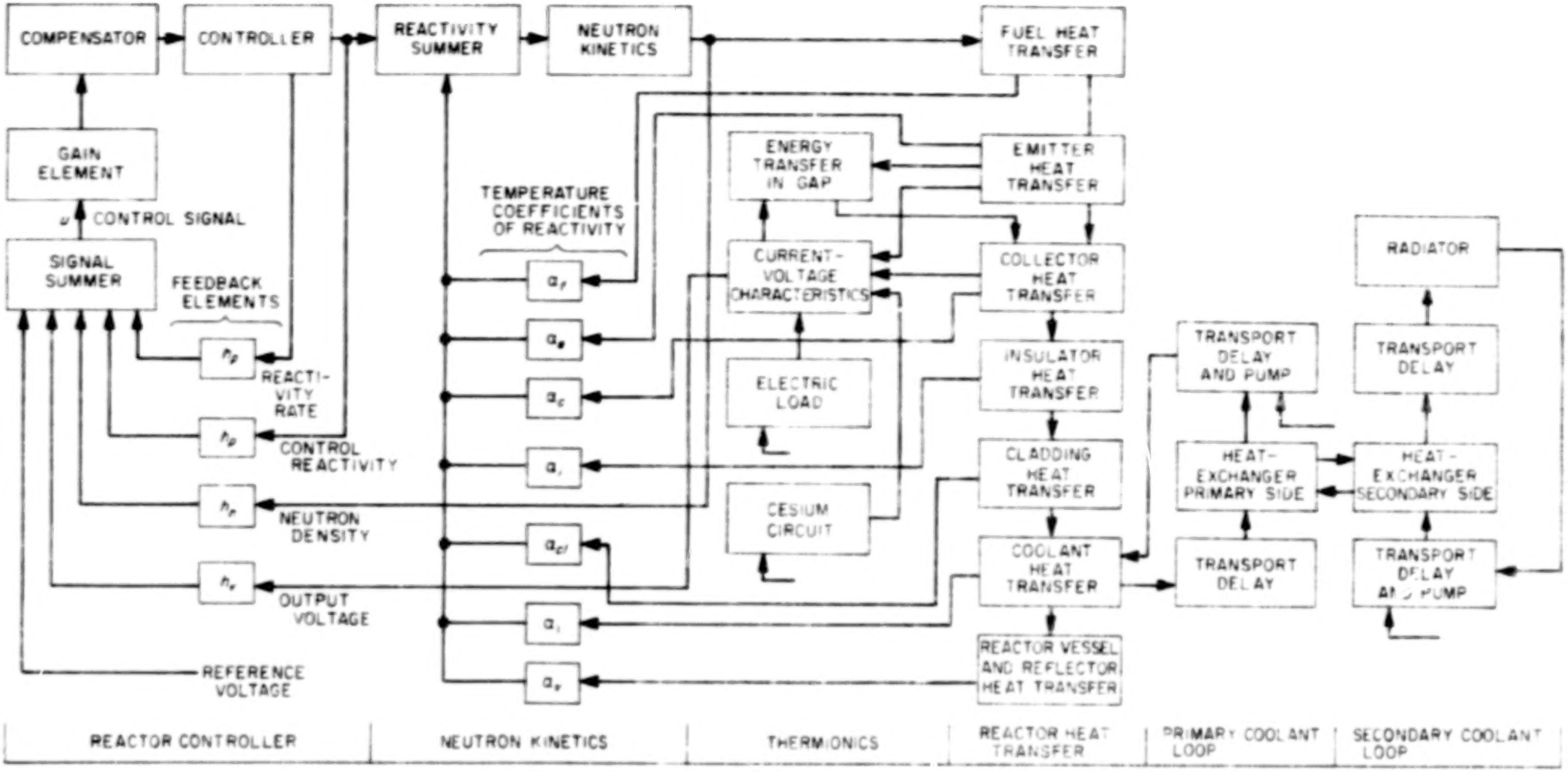

COMPENSATOR
CONTROLLER
REACTIVITY SUMMER
NEUTRON KINETICS
FUEL HEAT TRANSFER
GAIN ELEMENT
u CONTROL SIGNAL
SIGNAL SUMMER
FEEDBACK ELEMENTS
$h_{\dot\rho}$
REACTI-VITY RATE
h_ρ
CONTROL REACTIVITY
h_n
NEUTRON DENSITY
h_v
OUTPUT VOLTAGE
REFERENCE VOLTAGE
TEMPERATURE COEFFICIENTS OF REACTIVITY
α_f
α_e
α_c
α_i
α_{cl}
α_l
α_v
ENERGY TRANSFER IN GAP
CURRENT-VOLTAGE CHARACTERISTICS
ELECTRIC LOAD
CESIUM CIRCUIT
EMITTER HEAT TRANSFER
COLLECTOR HEAT TRANSFER
INSULATOR HEAT TRANSFER
CLADDING HEAT TRANSFER
COOLANT HEAT TRANSFER
REACTOR VESSEL AND REFLECTOR HEAT TRANSFER
TRANSPORT DELAY AND PUMP
HEAT-EXCHANGER PRIMARY SIDE
TRANSPORT DELAY
RADIATOR
TRANSPORT DELAY
HEAT-EXCHANGER SECONDARY SIDE
TRANSPORT DELAY AND PUMP
REACTOR CONTROLLER
NEUTRON KINETICS
THERMIONICS
REACTOR HEAT TRANSFER
PRIMARY COOLANT LOOP
SECONDARY COOLANT LOOP

Figure 3.- Simplified diagram for analysis of a thermionic reactor

$$\left.\begin{aligned}
\dot{x}_1 &= -\ 0.00164x_2 - 0.4x_3 - 0.00313x_4 + 10.0x_5 \\
\dot{x}_2 &= -\ 1.7x_2 + 319x_3 + 1.66x_4 \\
\dot{x}_3 &= 0.00031x_1 + 0.00035x_2 \\
&\qquad - 0.123x_3 - 0.00034x_4 + 0.0308x_5 \\
\dot{x}_4 &= +\ 0.0819x_2 - 0.113x_4
\end{aligned}\right\} \tag{1}$$

In this system of equations, x_1 is a measure of the change in precursor concentration, x_2 is the variation in collector temperature, x_3 is the change in output voltage, and x_4 is the variation in radiator temperature. The variable x_1 is related to the change in neutron density (δn, proportional to variation in thermal power) through the equation

$$\delta n = 10\rho + 0.1x_1 \tag{2}$$

where

$$\rho = x_5 + \rho_{fb}, \quad \text{total reactivity} \tag{3}$$

ρ_{fb} is the internal feedback reactivity and x_5 is the external reactivity.

Expression (1) is written with the voltage change x_3 as one of the state variables by elimination of the nuclear fuel-emitter temperature θ_{12}. The linearized relationships between the diode electrode temperatures and voltage and current δI are

$$x_3 = 0.00118\theta_{12} - 0.00020x_2 + 2.11\delta R_L. \tag{4}$$

$$\delta I = 0.0196\theta_{12} - 0.0036x_2 - 132\delta R_L \tag{5}$$

For the heat flow across the interelectrode gap, the relationship is

$$\delta q_{23} = 0.125\theta_{12} - 0.0357x_2 - 381\delta R_L \tag{6}$$

In Eqs. (4) through (6), the term arising from electric load perturbations, δR_L, has been retained to illustrate how it enters the system. In the analysis, however, only single input-single output systems are considered and, therefore the δR_L-terms are not included in expression (1). All inputs, except the reference voltage signal, are treated as system disturbances.

Figure 4 shows a Bode plot for the linear fourth-order, open-loop model and an experimentally determined (for small perturbations) Bode plot for the 19th order, non-linear mode. The fourth-order, open-loop transfer function is

$$G(s) = \frac{x_3(s)}{x_5(s)}$$

$$= \frac{0.031(s+1.79)(s+0.1)(s+0.033)}{(s+0.012 + j\ 0.016)(s+0.012 - j\ 0.016)(s+1.85)(s+0.066)} \tag{7}$$

THERMIONIC REACTOR CONTROL

Current-Voltage Characteristics

The unique feature of an in-core thermionic reactor is that the effect of electrical load perturbations is introduced directly and instantaneously into the reactor core. In conventional, non-boiling reactors the power is taken out via the coolant and the reactor tends to be inherently load-responsive due to a large overall negative temperature coefficient of reactivity. After the initial transient, the temperature drop across the reactor adjusts itself without control action to the new thermal power demand. In a thermionic reactor the response to electric load perturbations is leading the coolant, and since over half of the energy transport in the interelectrode gap is due to electron cooling, the effect is pronounced. (The net conversion efficiency, however, is generally about 15 percent in hardware diodes.)

Figure 5 shows calculated current-voltage curves for two cesium reservoir temperatures with the emitter temperature as a parameter and with the collector temperature held constant (ref. 12). For the lower reservoir temperature there is a cross-over of the characteristics, and the heat balance curve can pass through a maximum for a given loadline (for a constant voltage line). The higher emitter temperature condition is not stable, however. The emitter temperature increases and cesium desorbs from the emitter surface. This phenomenon, called "thermionic burn-out", reflects the fact that the cross-over is due to inadequate cesium pressure for the given thermal conditions.

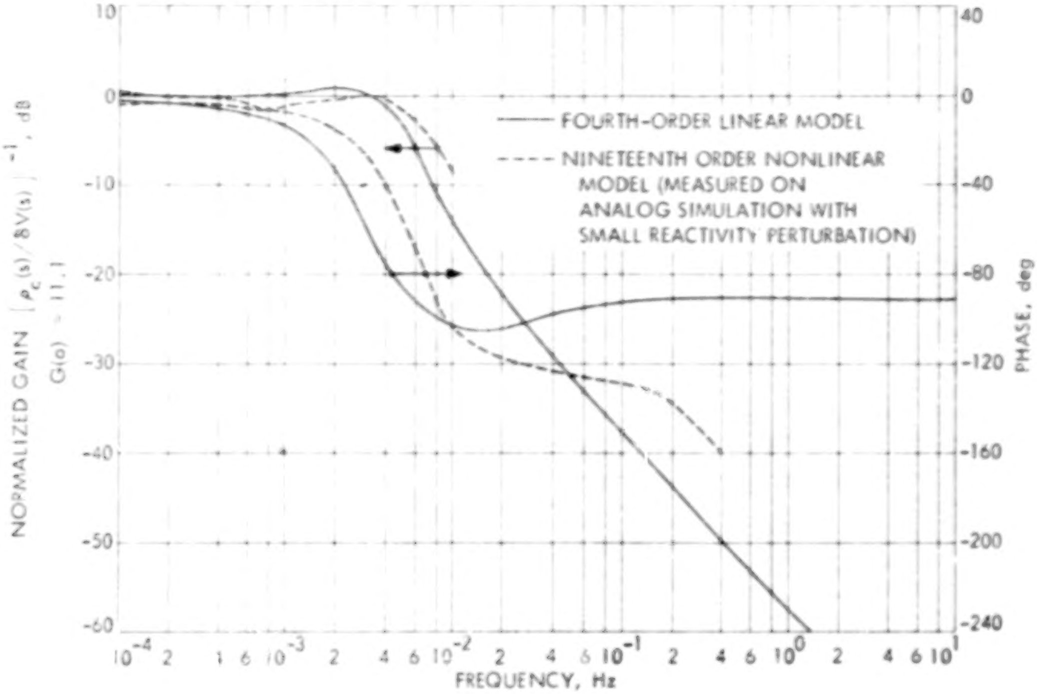


Figure 4.- Comparison of Bode plots for linear fourth order and non-linear open loop models

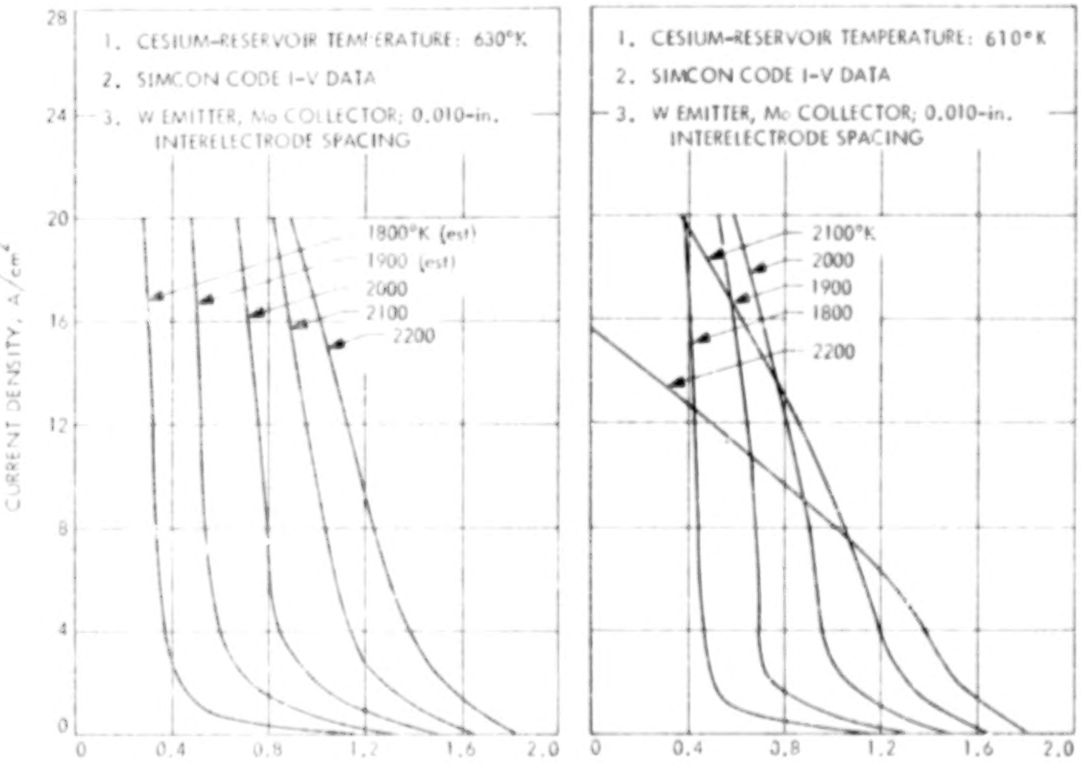


Figure 5.- Current density vs voltage characteristics at two cesium reservior temperatures

The cesium desorption point occurs at higher emitter temperatures and heat fluxes as the cesium reservoir temperature increases. If no reservoir temperature control were employed, the reservoir temperature would have to be set at higher than optimum at a noticeable loss in conversion efficiency (1 to 3 points) to provide an adequate margin of safety (ref. 4).

Method of Control

The control philosophy adopted for a thermionic reactor depends on the powerplant application. Past studies have shown that open-loop operation is unsatisfactory. Aside from programmed load changes, random load perturbations arise from component failures and subsequent switching in of standby modules, and from internal open and short-circuiting of thermionic diodes. If the electric load perturbation is relatively small and load compensating action is taken by the power conditioners within a few seconds, there is no need for adjustment of the reactor power level during the transient. The thermal lag of the reactor system prevents large temperature perturbations, even for relatively large, but short-term load variations. The operating point will initially follow a constant emitter temperature line on the current-voltage characteristics. Therefore a dead band and a dead time should be built into the controller, the levels of which are essentially set by the specifications on the output and the time it takes for the emitter to change temperature by a set amount. A stepping motor provides discrete reactivity insertions with the stepping rate limited by thermal and safety constraints. The details of the safety requirements are not known at this time, however.

For the purposes of the studies reported in this paper, dynamic characteristics, such as dead time, dead band, and saturation, are not considered in the analysis, although their effects are studied by analog computer simulation methods. In the analysis the controller is assumed to give a continuous signal to a continuously active control element driving motor. The controller design is derived from linear equations, and the control objective is to maintain a constant output voltage at the reactor terminals.

APPLICATION OF THE STATE VARIABLE FEEDBACK METHOD*

The State-Variable Feedback Design Technique

A linear system of differential equations may be described in matrix form by

---

*This section is a summary of work reported in reference 1.

$$\dot{\underline{x}} = A\underline{x} + \underline{b}u \tag{8}$$

$$y = \underline{c}^T\underline{x} \tag{9}$$

where $\underline{x}$ is an n-dimensional state vector, A is a constant nxn system matrix, b is an n-dimensional control vector, and u is the scalar control. The single output y is related to the state variables through the vector $\underline{c}$. The objective of state-variable feedback design is to apply feedback through constant gain and frequency-independent elements h_i, associated with their respective state variables x_i, so that a desired system transfer function is realized exactly. A representation in block diagram form is shown in Figure 6. The control signal is

$$u = r + \underline{h}^T\underline{x} \tag{10}$$

i.e., the sum of the reference input r and the feedback signals $h_i x_i$.

If some significant state variable is not available for measurement, it may be generated by a frequency-dependent feedback element, provided the mathematical relationship to a measurable state is known. In practice, only a few of the states are important for the feedback design, thus simplifying the controller.

A theorem due to Kalman (ref. 13) states that feeding back all states through linear constant-gain elements yields an optimal system with a quadratic performance index of the form

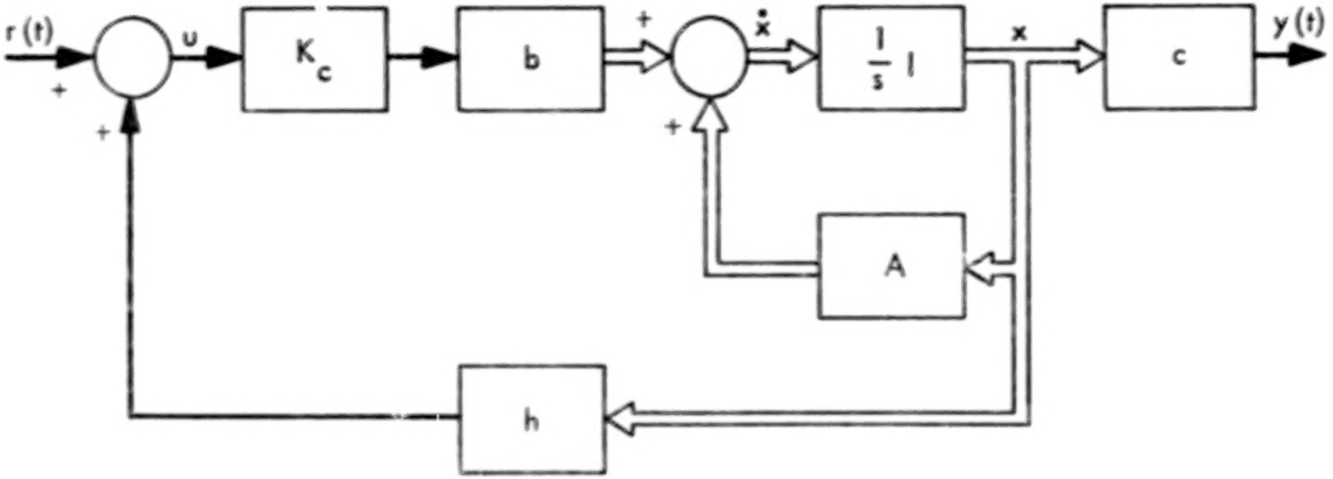
$$\text{P.I.} = \int_0^\infty (\underline{x}^T Q\underline{x} + \underline{u}^T P\underline{u})\, dt \tag{11}$$

In Eq. (11), Q must be at least positive semi-definite and symmetric, and P must be positive definite and symmetric. The state variable feedback method lends itself easily to digital computation (ref. 14).

If all system parameters are specified, except the feedback coefficients, a set of simultaneous linear algebraic equations in h_i are formed by equating the coefficients of powers of s in the denominator to like powers of s in the denominator of the desired system transfer function $[y(s)/r(s)]_d$. The one requirement is that the pole-zero excess of the desired transfer function must equal that of the given open-loop system transfer function.

THERMIONIC REACTOR CONTROL

STATE-VARIABLE FEEDBACK DESIGN TECHNIQUE



$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{b}u$$

$$y = \mathbf{c}^T\mathbf{x}$$

$$u = r + \mathbf{h}^T\mathbf{x}$$

$$\text{P.I.} = \int_0^{\infty}\left(\mathbf{x}^T Q\mathbf{x} + u^T P u\right) dt$$

Figure 6.- Block diagram of system with state variable feedback control

TABLE OF CONTENTS

1. SPACE VEHICLE CONTROL APPLICATIONS

Design Development of the Apollo Lunar Module
Kenneth J. Cox
Manned Spacecraft Center 1 1/A10

Design Development of the Apollo Command and Service Module Thrust Vector Attitude Control System
William H. Peters
Manned Spacecraft Center 63 1/E3

Orbiting Astronomical Observatory (OAO-A2) Stabilization and Control Subsystem
Thomas E. Huber
Goddard Space Flight Center 83 1/G10

Applications of the Hybrid Coordinate Method to the TOPS Autopilot
G. E. Fleischer
Jet Propulsion Laboratory 127 2/D7

Steering Laws for Double-Gimbal Control-Moment Gyros
Stephen W. Winder
Marshall Space Flight Center 157 2/F13

2. AIRCRAFT GUIDANCE, CONTROL AND HANDLING QUALITIES

Optimum Horizon Guidance Techniques for Aircraft
Heinz Erzberger and Homer Q. Lee
Ames Research Center 175 3/A8

Inertially-Aided, All-Weather Landing Systems for C/TOL.
Robert J. Pawlak
Electronics Research Center 185 3/B4

An Approach to the Determination of Aircraft Handling Qualities Using Pilot Transfer Functions
James J. Adams and Howard G. Hatch, Jr.
Langley Research Center 199 3/C4

A Frequency Domain Approach to Handling Qualities Design
William A. Wolovich
Electronics Research Center 217 3/D8

3. SYSTEM SIMULATION AND IDENTIFICATION

Optimal Regulator or Conventional? Setup Techniques for a Model Following Simulator Control System
Dwain A. Deets
Flight Research Center 237 3/F1

TABLE OF CONTENTS (CONT'D)

3. SYSTEM SIMULATION AND IDENTIFICATION (CONT'D)

Quasi-Optimal Control of a Moving-Base Simulator
Elwood C. Stewart
Ames Research Center 253 3/G3

Some Modern Control Techniques for Human Operator Modeling and Identification
Richard S. Shirley
Electronics Research Center 261 3/G11

4. ENGINE CONTROL AND ADVANCED PROPULSION TECHNIQUES

Normal Shock and Restart Controls for a Supersonic Airbreathing Propulsion System
George H. Neiner, Gary L. Cole, and Francis J. Paulovich
Lewis Research Center 299 4/D2

Optimal Control of a Supersonic Inlet to Minimize Frequency of Inlet Upstart
Bruce Lehtinen, John R. Zeller, and Lucille C. Geyser
Lewis Research Center 323 4/F1

Control System Studies for Thermionic Reactors
Richard J. Hermsen,
California State Polytechnic College
Henrik G. Gronroos
JPL, California Institute of Technology 337 4/G1

Application of Fluidics to New Control Components
Miles O. Dustin, Vernon D. Gebben, and Robert E. Wallhagen
Lewis Research Center 365 5/B7

Minimum Energy Control of a Class of Electrically Driven Vehicles
Yilmaz E. Sahinkaya
JPL, California Institute of Technology 387 5/D6

5. ADVANCED CONTROL TECHNIQUES

A New Formulation for the Epsilon Method Applied to the Minimum Time-to-Climb Problem
Lawrence W. Taylor, Jr., and Harriet J. Smith
Flight Research Center 423 5/G1

TABLE OF CONTENTS (CONT'D)

5. ADVANCED CONTROL TECHNIQUES (CONT'D)

Determination of Optimal Control Using Imbedding of the Terminal Conditions
Raymond C. Montgomery
Langley Research Center **435** 5/G13

Stability of Constant Gain Systems with Vector Feedback
George L. von Pragenau
Marshall Space Flight Center **443** 6/A10

Synthesis of Finite Settling Time Discrete Systems
James A. Gatlin
Goddard Space Flight Center **473** 6/C12

Design and Performance of Heart Assist or Artificial Heart Control Systems
John A. Webb, Jr., and Vernon D. Gebben
Lewis Research Center **495** 6/E1

System with Non-linear Controller

In a practical reactor controller, it is necessary to limit reactivity rate. It is important to ensure stability with this constraint; this can be achieved by constraining the control signal u, and not the reactivity rate itself as is shown in the section on Analog Simulation Results below.

A system with a non-linear element can be represented in block diagram form as shown in Figure 7. The equation $N = f(\sigma)/\sigma$ is a non-linear gain element with input σ and output $u = f(\sigma)$. The forward-loop transfer function is $G(s)$, with a single pole at $s = 0$ and with other poles in the left-half of the s-plane. The system equations are:

$$\begin{aligned} \dot{\underline{x}} &= A\underline{x} + \underline{b}u \\ u &= f(\sigma) \\ \sigma &= \underline{h}^T\underline{x} \end{aligned} \tag{12}$$

Popov's theory (ref. 6) gives a sufficient stability condition for the system of equations (Eq. 12). The Popov criterion applied to the diagram in Figure 7 states that the system is absolutely stable for any gain dependent-time invariant, single-valued non-linearity lying in the sector $0 \leq f(\sigma)/\sigma \leq k$ of the $[\sigma,\ f(\sigma)]$ - plane, if the following inequality is satisfied.

$$\mathrm{Re}\ [(1 + jq\omega)\ G(j\omega)\ H_{eq}(j\omega)] + \frac{1}{k} > 0,\ \omega \geq 0 \tag{13}$$

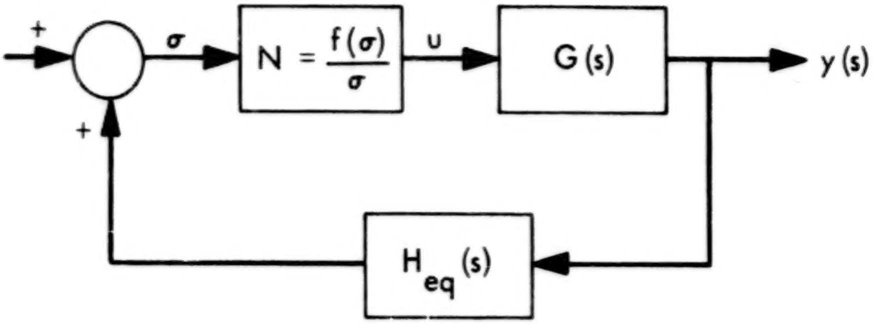
where Re denotes the real part of the function and q is a real number. If there is a pole at the origin, the stability sector does not include the σ-axis, but it can be arbitrarily close to this axis. A geometric formulation is obtained by plotting a modified frequency function defined by

$$W(j\omega) = \mathrm{Re}\ [G(j\omega)\ H_{eq}(j\omega)] - q\omega \mathrm{IM}\ [G(j\omega)\ H_{eq}(j\omega)] = x - qy$$

The stability criterion is now

$$x - qy + \frac{1}{k} > 0,\ \omega \geq 0 \tag{15}$$

The limiting condition is a straight line passing through the point $-1/k$ with a slop $1/q$ (the Popov line). The modified frequency function $W(j\omega)$ must be to the right of this line.



BLOCK DIAGRAM OF SYSTEM WITH NONLINEAR ELEMENT

$$\dot{\mathbf{x}} = A\mathbf{x} + b\mathbf{u}$$

$$\mathbf{u} = f(\sigma)$$

$$\sigma = \mathbf{h}^T\mathbf{x}$$

THE POPOV CRITERION

$$\mathrm{Re}(1 + jq\omega)\ G(j\omega)\ \mathbf{H}_{eq}(j\omega) + \frac{1}{k} > 0$$

$$W(j\omega) = \mathrm{Re}\left[G(j\omega)\mathbf{H}_{eq}(j\omega)\right] - q\omega \mathrm{Im}\left[G(j\omega)H_{eq}(j\omega)\right]$$

$$= x - qy$$

$$x - qy + \frac{1}{k} > 0$$

Figure 7.- Block diagram of system
with non-linear element

Controller Design

The reference thermionic reactor is controlled externally by actuators rotating drums in the reflector. The actuator transfer function is assumed to be

$$G_m(s) = \frac{2}{s^2 + 2s + 2} \tag{16}$$

This represents a system with a rise time of about 2 seconds and with a 5 percent overshoot.

The closed-loop response is taken to be the same as that of the controller, i.e., the desired system transfer function is chosen to be that of Eq. (16). Since the open-loop system has a pole-zero excess larger than 2, it is necessary to place the one extraneous pole so that its effect on the system dynamics is negligible.

Written in state-variable form, the amplifier-actuator equations are:

$$\dot{x}_5 = x_6$$

$$\dot{x}_6 = -\,2x_5 - 2x_6 + 2K_c u \tag{17}$$

To obtain the desired response function (Eq. 18) three zeros must be cancelled with three poles and the remaining pole is placed so that its contribution to the system dynamics is negligible. The closed-loop transfer function is therefore

$$\left[\frac{x_3(s)}{r(s)}\right]_d = \frac{200\ (s+1.79)(s+0.1)(s+0.0329)}{(s+1+j)(s+1-j)(s+100)(s+1.79)(s+0.1)(s+0.0329)} \tag{18}$$

The computer calculated values for the feedback elements h_i and the controller gain K_c required to obtain the desired system transfer function are given below.

$h_1 = 2.92 \times 10^{-4}$	precursor concentration
$h_2 = 1.53 \times 10^{-5}$	collector temperature
$h_3 = 0.939$	voltage
$h_4 = 8.03 \times 10^{-6}$	radiator temperature

$h_5 = 3.05 \times 10^{-2}$ control reactivity

$h_6 = -1.54 \times 10^{-2}$ control reactivity rate

$K_c = 3250$ gain constant

Analog Simulation Results

A multitude of analog computer simulation runs were made to evaluate the performance of the controller designed by the state variable feedback method. Both the linear powerplant model given by Eq. (1) and the non-linear model described in reference 1 were used. A few selected results are discussed below

Figure 8 compares the response of the linear and non-linear systems to a + 0.035-volt step demand in voltage. For the linear system, the output voltage trajectory follows the desired transfer function. The higher order dynamics of the non-linear model manifests itself with a few oscillations before equilibrium is reached. The differences between the trajectories are minor, however. Note the large reactivities and reactivity rates.

Figure 9 compares response trajectories for a -0.02 Ω-cm^2 step change in electric load. This is not, as mentioned earlier, a simple input disturbance, and there is a small output voltage steady-state error. The curves in Figure 9 again show very large reactivities and reactivity rates. Also note the "spike" in the output voltage, which must be eliminated by the power conditioning equipment.

There are no significant differences between the linear and non-linear models in their response to loss of a feedback element. Loss of reactivity feedback h_5 causes an instability, and loss of reactivity rate h_6 causes oscillations to occur. Loss of voltage feedback h_3 causes the output voltage to increase to a new, much higher level. A limit on the control signal eliminates instabilities from loss of h_6. This is not true if the reactivity or reactivity rate are limited. Elimination of the feedbacks from collector temperature and radiator temperature, has no noticeable effect on the response trajectories.

The control signal u is constrained such that the reactivity rate does not exceed some set value. To investigate the effects of this on system stability, the Popov criterion described earlier is applied. Figure 10 is a plot of the modified frequency function (Eq. 15) for the system. It is apparent from the figure that the Popov line can be placed arbitrarily close to the origin; therefore limiting u causes no stability problem. With the control signal constrained, loss of reactivity feedback h_5 causes the system to oscillate with a small amplitude. The oscillations can

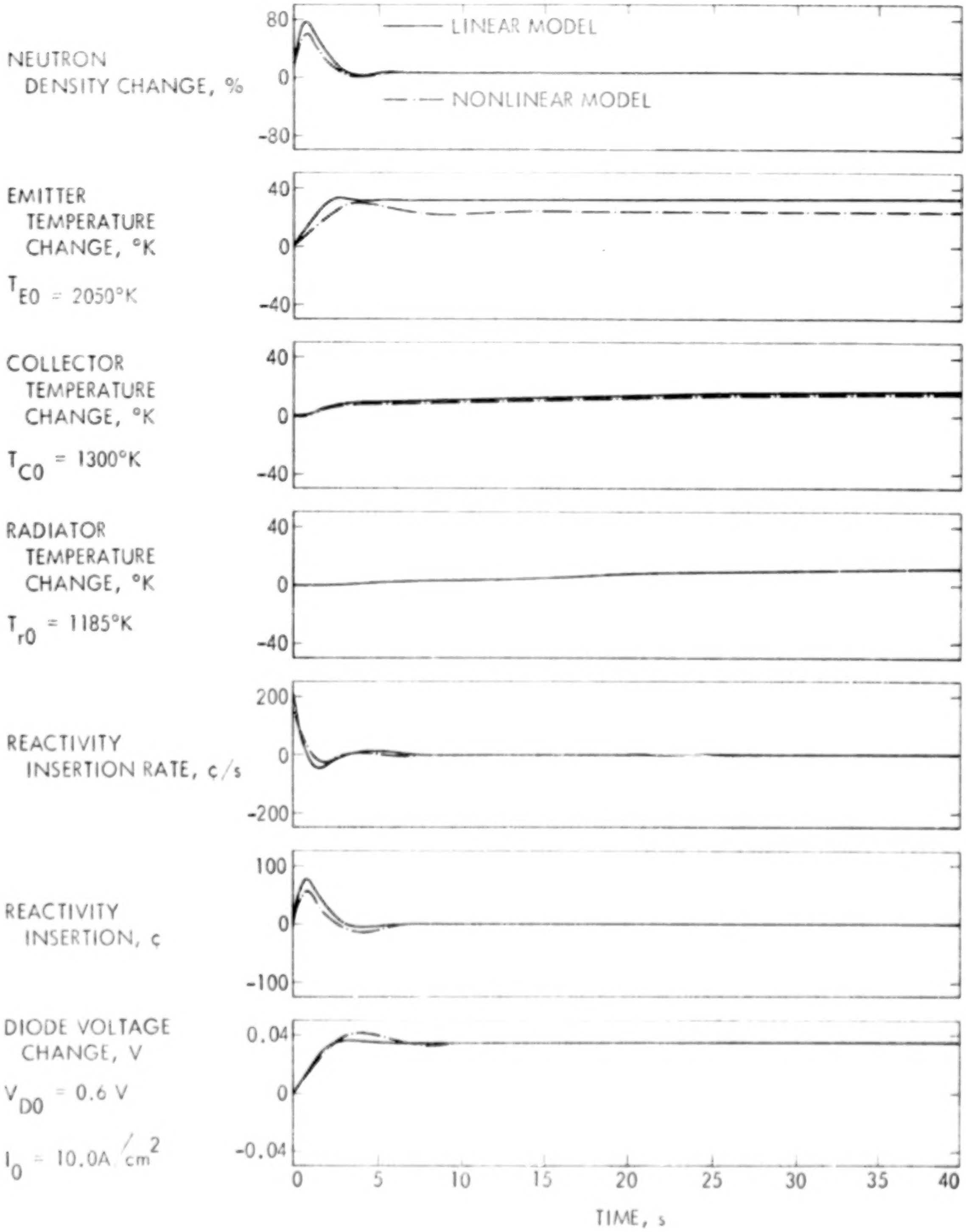


Figure 8.- Response of the linear and non-linear system models to a + 0.035 V step in reference voltage (no control signal constraint

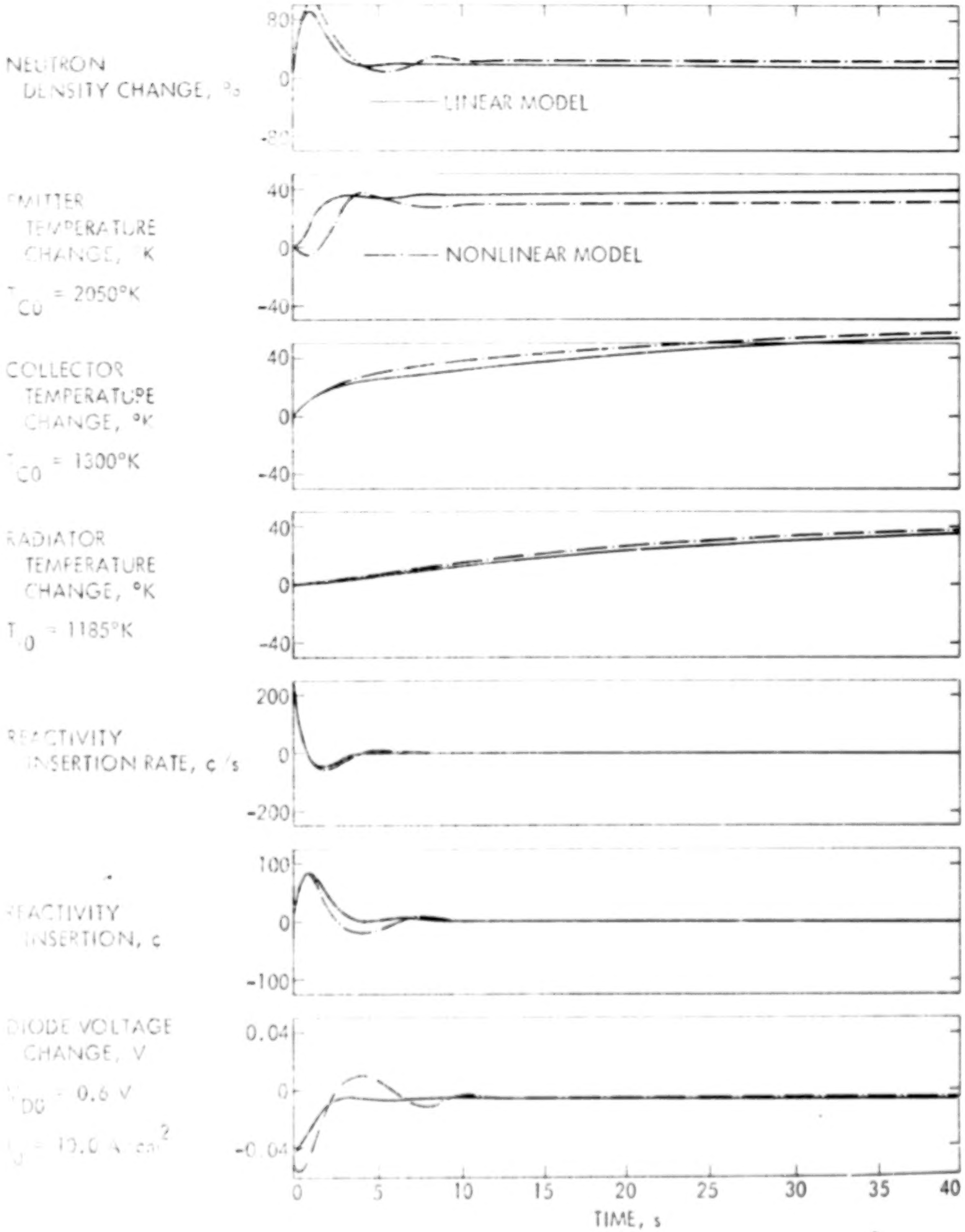


Figure 9.- Response of the linear and non-linear system models to a - 0.02 Ω cm^{-2} step in electric load (no control signal constraint)

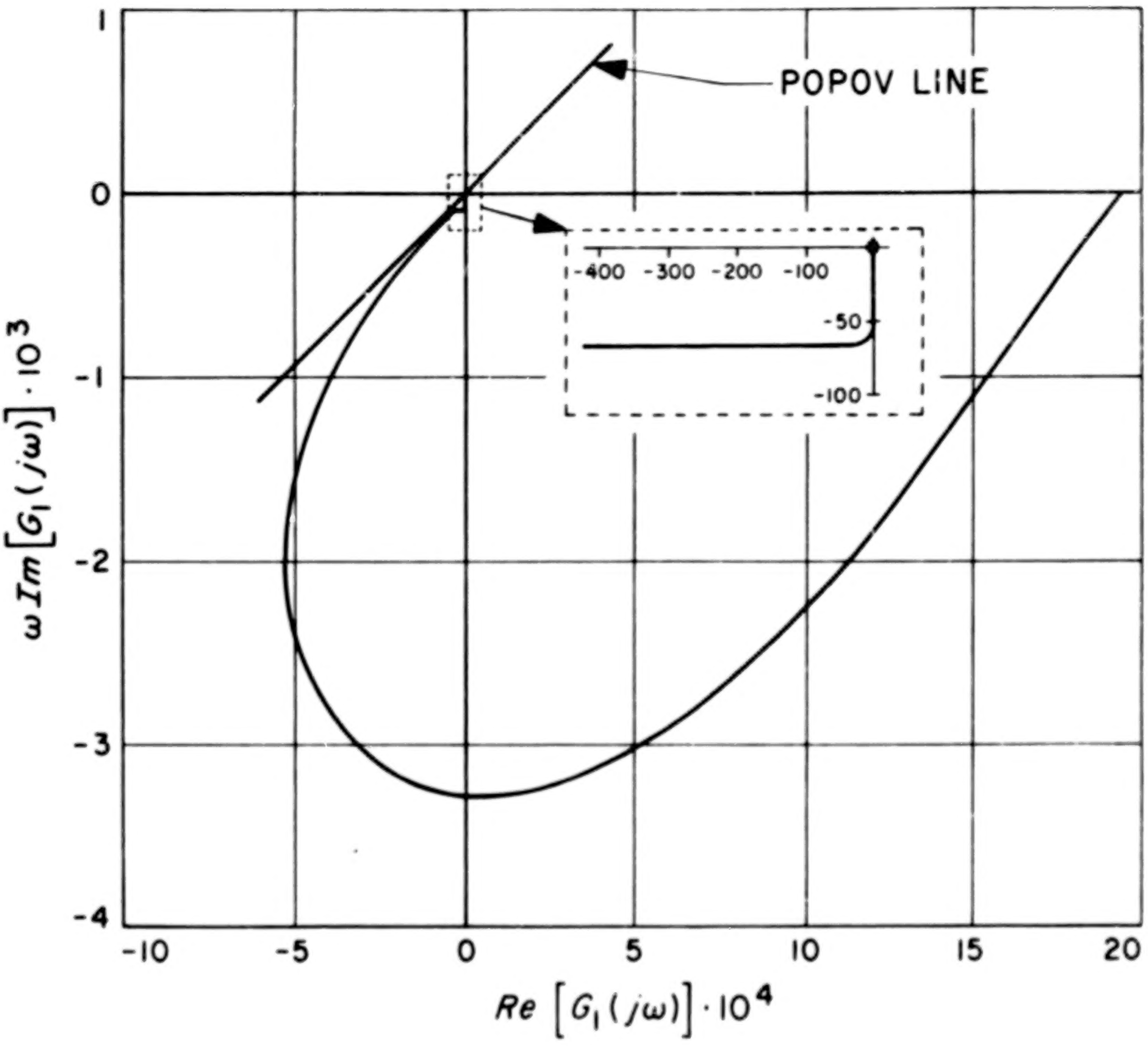


Figure 10.- Modified frequency function for controlled system

be made to disappear effectively by increasing the gain by a factor of 3. Figure 11 illustrates the response of the high-gain linear system to a +0.035-volt step in reference voltage when the control signal is constrained to limit the reactivity rate to 20¢/sec.

APPLICATION OF THE ANALYTICAL DESIGN METHOD

Preliminary Considerations

Although the state variable feedback method does identify the important states and does give insight into the control problem, the above noted shortcomings prompted the application of the analytical design method (ref. 7). This method allows incorporation of constraints on a state, and also in an approximate manner incorporation of criteria for response time, peak overshoot, and steady state error. Furthermore, for simplified system models, hand calculations are feasible.

Considering the inherently slow responding plant and reconsidering system specifications, a response time of 8 seconds is deemed adequate.

The analytical design method permits use of simplified models obtained directly from measured or calculated Bode plots of a more complete powerplant model. Thus the Bode plot of the high-order thermionic reactor powerplant model in Figure 4 is approximated by

$$G_R(s) = \frac{v(s)}{\dot{\rho}(s)} = \frac{3.61\ (s/0.0755 + 1)}{s\left[\left(\frac{s}{0.0377}\right)^2 + \frac{1.41}{0.0377}\,s + 1\right]\left(\frac{s}{0.314} + 1\right)} \qquad (19)$$

Since this is a feasibility study, it is desirable to consider the thermionic reactor system independently of a specified control motor. A motor-tachometer is assumed with the transfer function

$$\frac{\rho(s)}{v_m(s)} = \frac{1}{s\ (\tau s + 1)} \qquad (20)$$

where τ is sufficiently small to be neglected.

The Analytical Design Method

The basic block diagram of the system under study is shown in Figure 12. The objective is to find a compensating transfer function $G_C(s)$ subject to the performance index

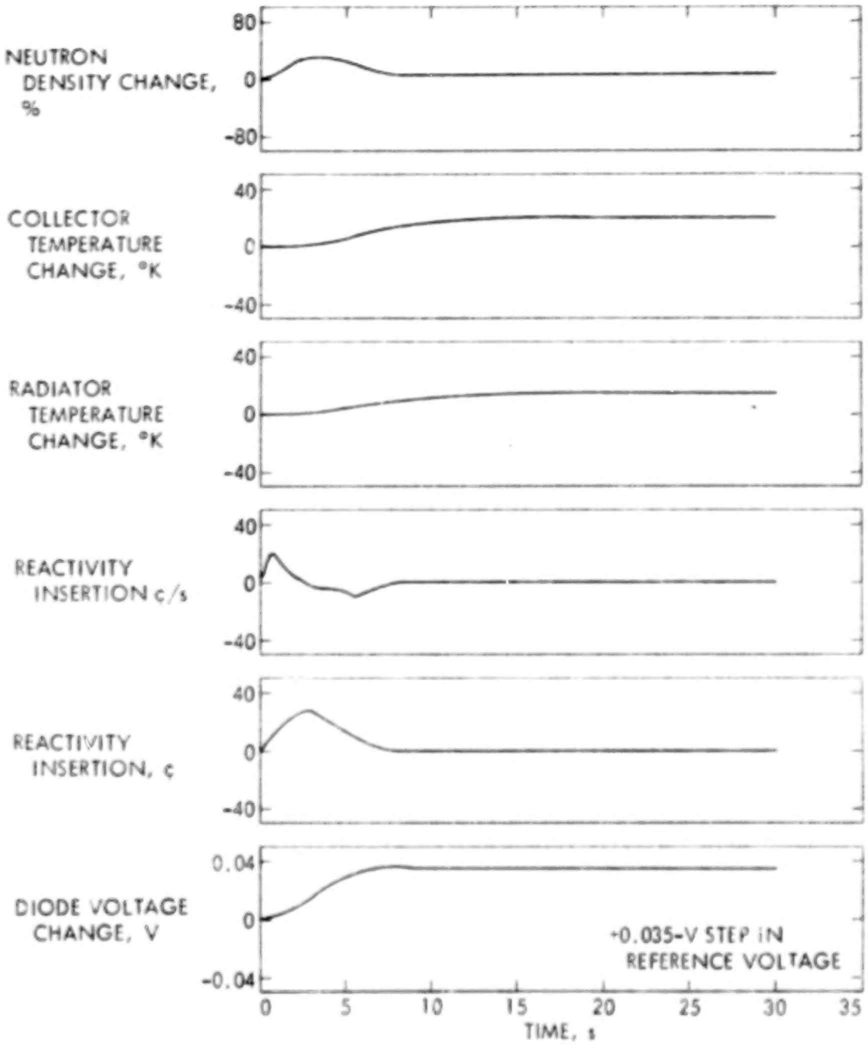
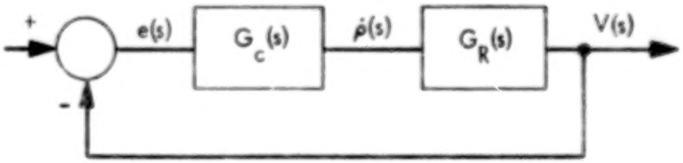


Figure 11.- Response of linear system model to a + 0.035 V step in reference voltage (final controller design with a control signal constraint to limit reactivity rate to 20¢/sec)

THERMIONIC REACTOR CONTROL

APPLICATION OF ANALYTICAL DESIGN METHOD



$$\text{P.I.} = \int_0^{\infty} \left[e^2(t) + k^2 \dot{\rho}^2(t) \right] dt$$

THE OPTIMAL SYSTEM TRANSFER FUNCTION

$$F_0 = (1/YZ)\ (Z/\overline{Y})$$

Figure 12.- Application of the analytical design method

$$P.I. = \int_0^\infty [e^2(t) + k^2 \dot{\rho}^2 (t)]\, dt \qquad (21)$$

i.e., the time integral over the sum of the error squared and the weighted reactivity rate squared. The theory in reference 7 derives the optimal system transfer function F_o for the performance index in Eq. (21) to be

$$F_o = (1/YZ)(Z/\sqrt{\bar{Y}})_+ \qquad (22)$$

where $(Z\sqrt{\bar{Y}})_+$ denotes the term stemming from the poles of $Z\sqrt{\bar{Y}}$ in the left-half of the s-plane.

In Eq. (22):

$$Y = [1 + k^2/G_R\bar{G}_R]^+ \qquad (23)$$

(all poles and zeros of the entire term are in the left half of the s-plane),

$$\bar{Y} = [1 + k^2/G_R\bar{G}_R] \qquad (24)$$

(all poles and zeros of the entire term are in the right half of the s-plane),

$$Z = [V_d\bar{V}_d]^+ \qquad (25)$$

(all poles and zeros of the entire term are in the left half of the s-plane).

and $\bar{V}_d(s) = V_d(-s)$

$$\bar{G}_R(s) = G_R(-s) \qquad (26)$$

Synthesis of the Series Compensation

In the design of the thermionic reactor three broad criteria are of concern:

1. The steady-state error to a command change in voltage or to a load change should be less than the dead zone limits.
2. The system must be absolutely stable.
3. The transient response must satisfy specifications with respect to response time and peak overshoot.

Limiting the integral of the squared error, $e^2(t) = [v_d(t) - v(t)]^2$ to less than a specified constant satisfies the above requirements. The constraint on the reactivity rate $\dot{\rho}$ is incorporated into the control system design by minimizing the integral of the square of the reactivity rate. The weighting factor k^2 in the performance index of Eq. (21) indicates the relative importance between the system error and reactivity rate input constraint. The procedure used here is to minimize the integral of the square of the reactivity rate subject to the values of the integral of the error squared as determined approximately by the response time requirement.

For a step input the optimal closed-loop transfer function [Eq. (22)] becomes:

$$F_o(s) = (s/Y)\ [1/s\ Y(o)] \qquad (27)$$

Substituting Eq. (19) for $G_R(s)$ into the product of Eqs. (23) and (24) gives:

$$Y\bar{Y} = 1 + \frac{k^2(-s^2)\left[\left(\frac{s}{0.0377}\right)^2 + \frac{1.41}{0.0377}s + 1\right] x \left(\frac{s}{0.314} + 1\right)\left[\left(\frac{-s}{0.0377}\right)^2 + \frac{1.41}{0.0377}(-s) + 1\right]\left(\frac{-s}{0.314} + 1\right)}{3.61^2\left(\frac{s}{0.0755} + 1\right)\left(\frac{-s}{0.0755} + 1\right)} \qquad (28)$$

From Eq. (28) $Y(0) = 1$ and Eq. (27) reduces to

$$F_o(s) = 1/Y(s) \qquad (29)$$

The zeroes of $F_o(s)$ are the poles of $Y(s)$ and are obtained by inspection of Eq. (28). The poles of $F_o(s)$ are the zeroes of $Y(s)$ and are to be determined. These zeroes are found by inverting Eq. (28) and forming the root locus plot of the resulting denominator in the $(-s^2)$-plane. Rather than selecting a value for k^2 to determine the roots, the response time requirement (8 seconds) is used by applying the approximate rule

$$t_R = \pi/d \tag{30}$$

where t_R is the time to 90 percent of final value, and d is the distance from the origin to the nearest root on the $(-s^2)$-plane. Neglecting the dipoles the optimal closed-loop transfer function as determined from the root locus plot becomes

$$F_o(s) = \frac{1}{\left(\frac{s}{0.456} + 1\right)\left[\left(\frac{s}{0.392}\right)^2 + \frac{1.12}{0.392}\, s + 1\right]} \tag{31}$$

The corresponding open-loop transfer function is

$$G_o(s)^{-1} = F_o(s)^{-1} - 1 = [G_R(s)\ G_c(s)]^{-1} \tag{32}$$

Inserting $G_R(s)$ from Eq. (19) and $F_o(s)$ from Eq. (31) into Eq. (32) gives:

$$G_c(s) = \frac{0.0544\left[\left(\frac{s}{0.0377}\right)^2 + \frac{1.41}{0.0377}\, s + 1\right]\left[\frac{s}{0.314} + 1\right]}{\left(\frac{s}{0.0755} + 1\right)\left[\left(\frac{s}{0.598}\right)^2 + \frac{1.51}{0.598}\, s + 1\right]} \tag{33}$$

$G_c(s)$, as given by Eq. (33), is relatively complex but can be approximated by

$$G_{ca}(s) = \frac{0.0544\left[\frac{s}{2\pi \cdot 10^{-2}} + 1\right]^2}{\left[\left(\frac{s}{16\pi \cdot 10^{-2}}\right)^2 + \frac{1.8}{16\pi \cdot 10^{-2}}\, s + 1\right]} \tag{34}$$

Analog Simulation Results

The obtained compensator [Eq. (33)] was programmed on an analog computer, together with the full non-linear model for the reference thermionic reactor powerplant. Figure 13 shows the response obtained with the analytical design method for a +0.035-volt step. The approximate compensator [Eq. (34)] was also programmed and the results obtained are identical. Note the low values of the reactivity rate. The response trajectories obtained with the analytical design method cannot be compared directly with

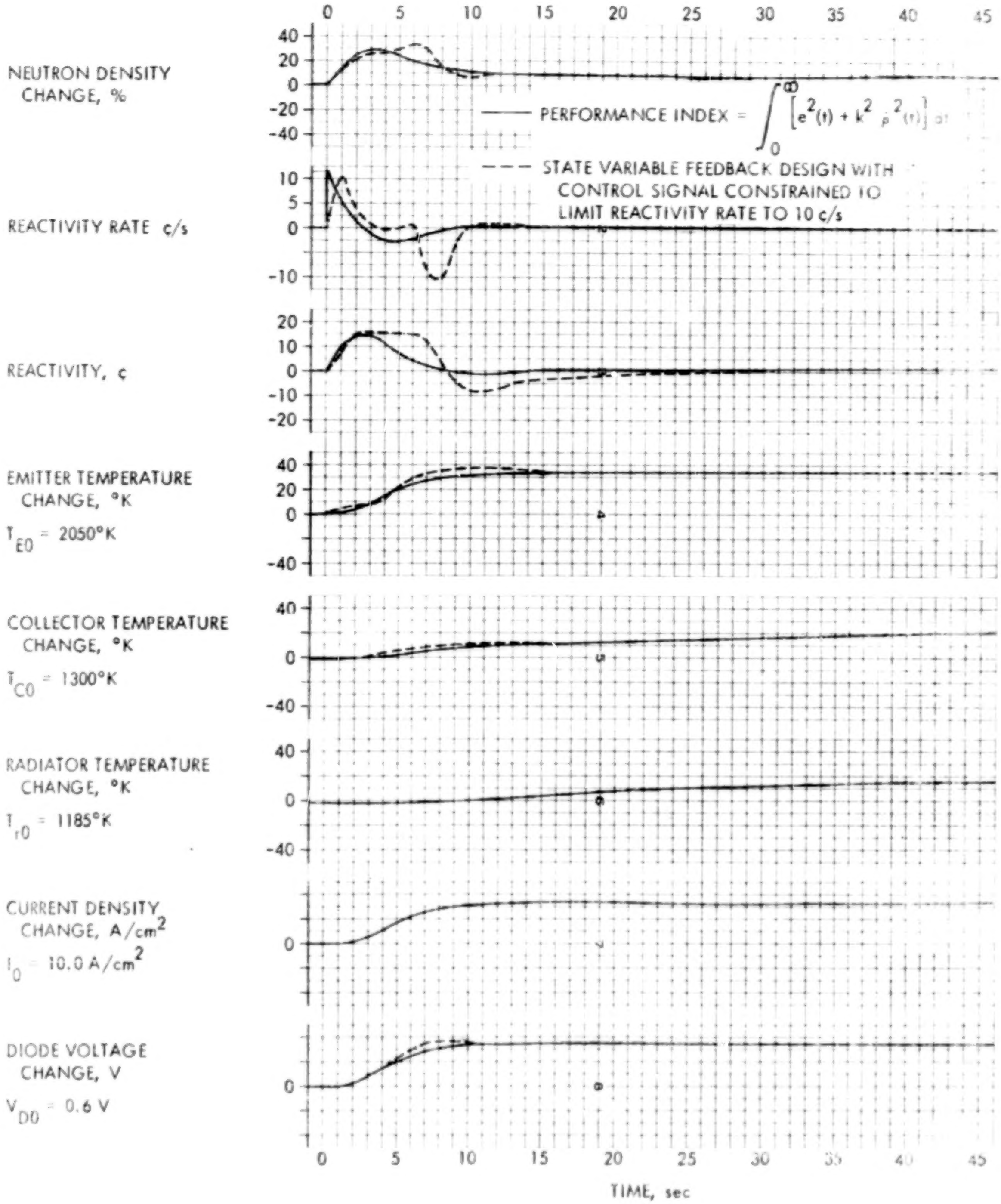


Figure 13.- Response of non-linear system to a +0.035 V step in reference voltage with controller designed by analytical design method

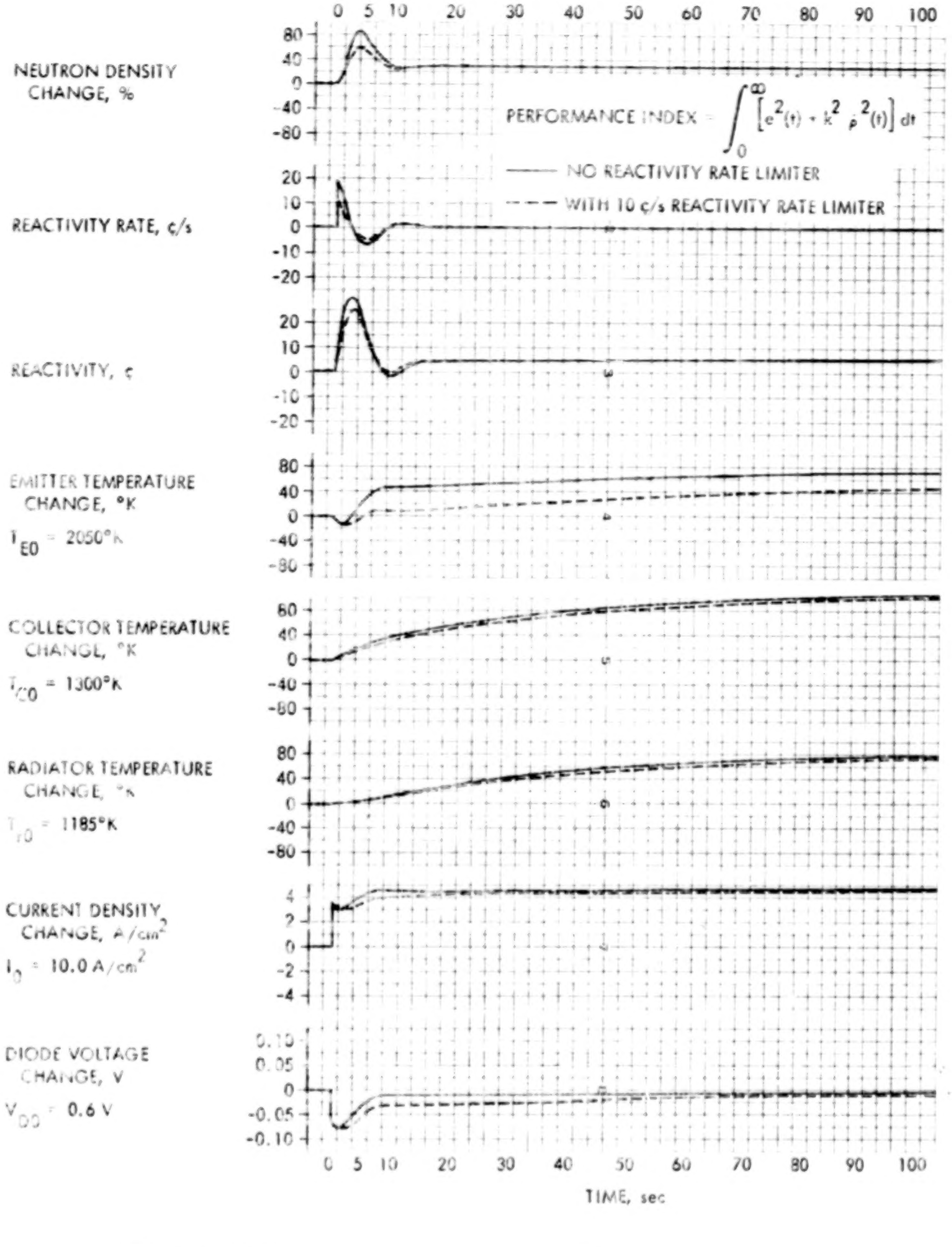


Figure 14.- Response of non-linear system to a $-0.02\ \Omega\ \mathrm{cm}^{-2}$ step in electric load with controller designed by analytical design method (with and without 10¢/sec reactivity rate limit

those obtained with the state variable feedback method because of different response time criteria. For illustration the trajectories that resulted from the state variable feedback analysis with a control signal constraint ($\dot{\rho} < 10$ ¢/sec) are also indicated in Figure 13.

Figure 14 shows the responses for a $-0.02\ \Omega\text{-cm}^2$ step change in electric load with and without a 10 ¢/sec limiter. By applying the Popov criterion [Eq. (15)] it is shown that no stability problems result from constraining the reactivity rate.

REFERENCES

1. Gronroos, H.: "Analog Studies of Thermionic Reactor Dynamics." JPL SPS 37-45, vol. IV, June 30, 1967.
2. Shapiro, J. L., and Gronroos, H.: "Stability Analysis of a Thermionic Reactor." Presented at the 1969 Winter Meeting of the American Nuclear Society, San Francisco, Calif., Nov. 30 - Dec. 4, 1969.
3. Gronroos, H., and Davis, J. P.: "Thermionic Diode Kinetics Experiment - Operation and Results." IEEE Conference Record of the 1969 Thermionic Conversion Specialist Conference, Carmel, Calif., Oct. 21-23, 1969.
4. Hermsen, R. J., and Gronroos, H.: "Study of Cesium Reservoir Control for Thermionic Reactors." JPL SPS 37-59, Vol. III, Oct. 31, 1969.
5. Schock, A.: "Effects of Cesium Pressure on Thermionic Stability. IEEE Conference Record of 1968 Thermionic Conversion Specialist Conference, Framingham, Mass., Oct. 21-23, 1968 (68 C 58-ED, pp. 313-317).
6. Weaver, L. E.: Reactor Dynamics and Control. American Elsevier, 1968.
7. Chang, S. S. L.: Optimum Control Systems. McGraw-Hill, 1961.
8. "Thermionic Reactor Power Plant Design Study." General Electric NTPO, Report No. GESR-2115, Vol. I and II, July 1968 (CRD).
9. Papers in Sessions 4, 9 and 29 in Proceedings of the Fourth Intersociety Energy Conversion Engineering Conference, Washington, D. C., Sept. 22-26, 1969. American Institute of Chemical Engineers, New York, N. Y.
10. Davis, J. P.: "Review of the Status of Thermionic Reactor Systems." Paper No. 69-WA/Ener-5, ASME Winter Annual Meeting, Los Angeles, Calif., Nov. 16-20, 1969.

11. Weaver, L. E., Gronroos, H. G., Guppy, J. G., and Davis, J.P.: "A Control System Study for an In-Core Thermionic Reactor." JPL Technical Report 32-1355, Jan. 15, 1969.

12. Wilkins, D. R.: "SIMCON - A Digital Computer Program for Computing Thermionic Converter Performance Characteristics." General Electric NTPO, Report No. GESR-2109, Feb. 1, 1968.

13. Kalman, R. E.: "When Is a Linear System Optimal? " J. Basic Eng., ser. D, vol. 66, March 1964, pp. 51-60.

14. Melsa, J. L.: "A Digital Computer Program for the Analysis and Design of State Variable Feedback Systems." Engineering Experiment Station Report, University of Arizona, College of Engineering, Tucson, Ariz., March 1967.

APPLICATION OF FLUIDICS TO NEW CONTROL COMPONENTS

By Miles O. Dustin, Vernon D. Gebben, and Robert E. Wallhagen
Lewis Research Center

SUMMARY

The inherent tolerance for nuclear radiation makes fluidic devices candidates for nuclear rocket control systems. Also, they are being considered for supersonic jet engine inlet control because of their high temperature and vibration tolerance. This paper describes three new control components being considered for these applications.

A fluidic circuit to control a pneumatic stepping motor for nuclear rocket control drum actuation is discussed. The circuit concepts were first evaluated in a breadboard model utilizing commercial amplifiers to a great extent. A completely integrated circuit was then designed, built, and evaluated. The motor has a nutating gear driven by eight bellows. The nutating gear drives an output gear which advances 0.25 degree for each input pulse. The fluidic driving circuit is a ring-type counter which furnishes pressures to the bellows in response to input pressure pulses.

An all-fluidic sensor is being developed for determining the position of the normal shock in the inlet of a supersonic jet engine. The sensor was evaluated in wind tunnel tests on a Mach-2.5 mixed compression inlet. The sensor uses fluid jet amplifiers to determine shock position from the shape of the wall static pressure profile. Both static and dynamic tests were performed in the wind tunnel.

A new vortex valve configuration has been developed to prevent supersonic jet engine inlet unstarts by regulating bypass flow. The vortex valves are controlled by the pressure gradient across the normal shock. When the shock is downstream of the valve, a strong vortex is induced in the valve which blocks bypass flow. If the shock moves upstream of the valve, the pressure gradient is reduced, thereby reducing the vortex and permitting high bypass flow rate. The special vortex valve has a 10 to 1 throttling range and a high grain. These characteristics were obtained by the irregular shape of the peripheral wall of the valve.

INTRODUCTION

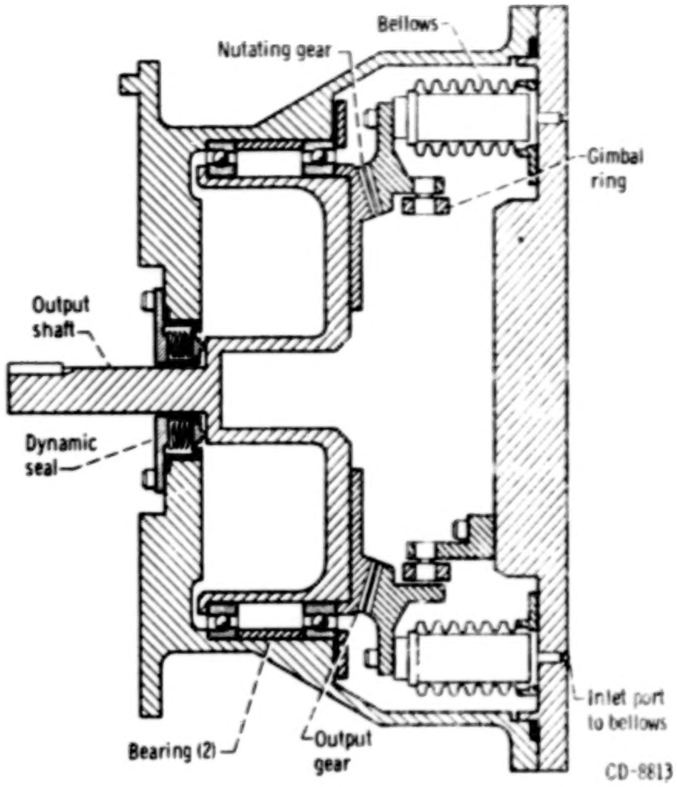
Fluid amplifiers have many inherent characteristics which make them appear well suited to certain aerospace applications. They can be made of radiation-resistant materials, and their operation appears to be inherently radiation-insensitive. They are, therefore, being considered for control circuits which must be located near the nuclear rocket engine. Since they can be made from materials that withstand high temperatures, and since they can be powered directly from the engine air, they also look attractive for jet engine control systems. Since there are no moving parts, their wear rate should be low and their reliability high.

This report describes three control components which use fluidic devices. The first system was intended for use on a nuclear rocket engine, and the other two for supersonic jet engine inlet applications.

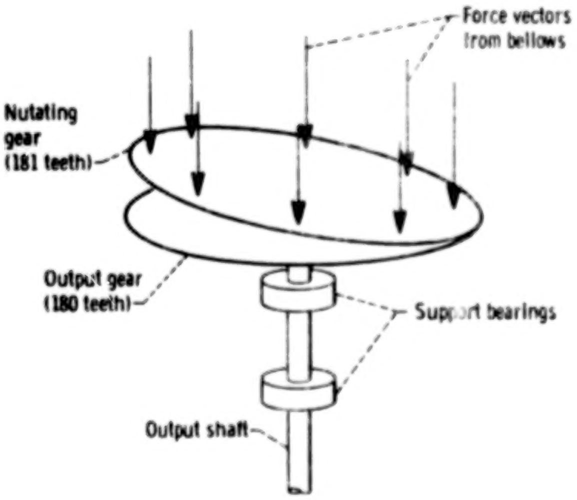
PNEUMATIC STEPPING MOTOR ACTUATOR

The nuclear rocket power level is controlled by rotating a number of drums located in the reactor shell. These drums are loaded with neutron-absorbing material. High pressure hydrogen is available for operation of the control drum actuators.

A pneumatic stepping motor (ref. 1) designed for potential use as a control drum actuator is shown in Figure 1. The two basic parts of the stepping motor are an output gear, which is free to rotate only, and a nutating gear, which is free to nutate or wobble only. The nutating gear is actuated by eight bellows located on the periphery of the gear. When four adjacent bellows are pressurized, the gear is tilted and forced into contact with the output gear. Sequencing the bellows pressurization pattern, as shown in Figure 2, causes the nutating gear to nutate. It is important to note here that during each step when a new bellows is pressurized the opposite bellows is depressurized. For example, to advance from the pulse 1 state to the pulse 2 state, $\overline{B}$ bellows are pressurized and the B bellows are depressurized. The output gear has 180 teeth and the nutating gear 181. This means that for each complete cycle of the nutating gear, the output gear must advance one tooth, or two degrees of rotation. Since there are eight bellows, each step increment of the bellows pressurization pattern causes the output gear to rotate 0.25 degree.


Bellows
Nutating gear
Gimbal ring
Output shaft
Dynamic seal
Inlet port to bellows
Bearing (2)
Output gear
CD-8813

(a) Cross-sectional view.


Force vectors from bellows
Nutating gear (181 teeth)
Output gear (180 teeth)
Support bearings
Output shaft

(b) Schematic representation.

Figure 1.- Pneumatic stepping-motor actuator

A digital fluidic circuit was developed to control the stepping motor actuator (ref. 2). The circuit was breadboarded with conventional fluidic components. After it was determined that the breadboard circuit could meet the system requirements, a complete integrated circuit was designed and built. The integrated circuit is described in reference 3.

The principal portion of the fluidic circuit is the counting circuit shown in Figure 3. This circuit controls the bellows pressurization pattern. The counting circuit stores the pattern and rotates it either forward or backward in response to command pulses. The outputs are labeled A, $\bar{A}$,to D, $\bar{D}$ and correspond to the actuator bellows positions indicated in Figure 2. Forward command pulses enter at eight locations marked T_f. Likewise, the backward pulse enters at eight locations marked T_b. To advance the bellows pressurization pattern by one step, only one of the central bistable amplifiers should switch. The passive AND elements act as permissive elements that only allow the correct central bistable amplifier to switch. For example, assume that outputs, B, C, D, and $\bar{A}$ are in the "on" condition when a forward command pulse enters the circuit. In Figure 3 the darker lines indicate those signals that are in the "on" condition. Amplifier B should be the only amplifier to switch, and it can switch only if amplifier A has already switched to its $\bar{A}$ leg. The passive AND element number 11 will have its supply port pressurized, thus allowing the pulse to pass, and switch central bistable amplifier B.

For proper operation, it is important that the input command pulses are less than 1 millisecond in duration. Delays marked τ in Figure 3 allow the input pulses to be complete before the counting circuit is set for the next input pulse.

Two pulse-conditioning circuits perform the desired pulse shaping. One of these circuits, shown in Figure 4, is required for the forward command pulses and one is required for the backward command pulses. Amplifier 1, which is an OR-NOR element, converts the command pulse to a push-pull signal of fixed amplitude. The pulse width is fixed by amplifier 3, a bistable element. The pulse from amplifier 2 is applied simultaneously to an orifice feeding control port C_1 and to a millisecond delay time which feeds control port C_2. Thus, the pulse initially causes a pressure at control port C_1 which switches amplifier 3 to its R_2 leg. One millisecond later the pulse reaches control port C_2 of amplifier 3 by way of the delay line. Amplifier 3 is now switched back to its R_1 leg, thus terminating the 1-millisecond pulse. Amplifier 4 serves as a power amplifier to provide sufficient flow and pressure to feed the eight input ports of the counting circuit.

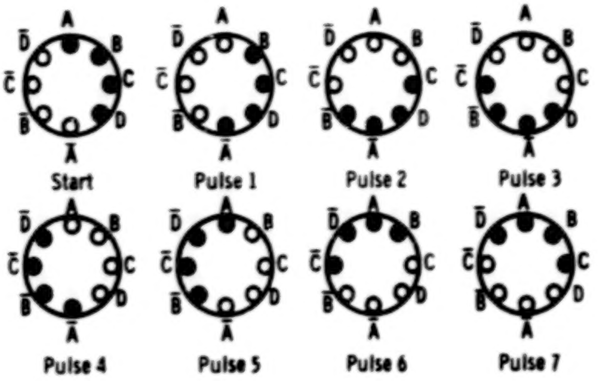

A
B
C
D
Ā
B̄
C̄
D̄
Start
Pulse 1
Pulse 2
Pulse 3
Pulse 4
Pulse 5
Pulse 6
Pulse 7

Figure 2.- Sequencing of bellows pressurization pattern by forward-counting input pulses

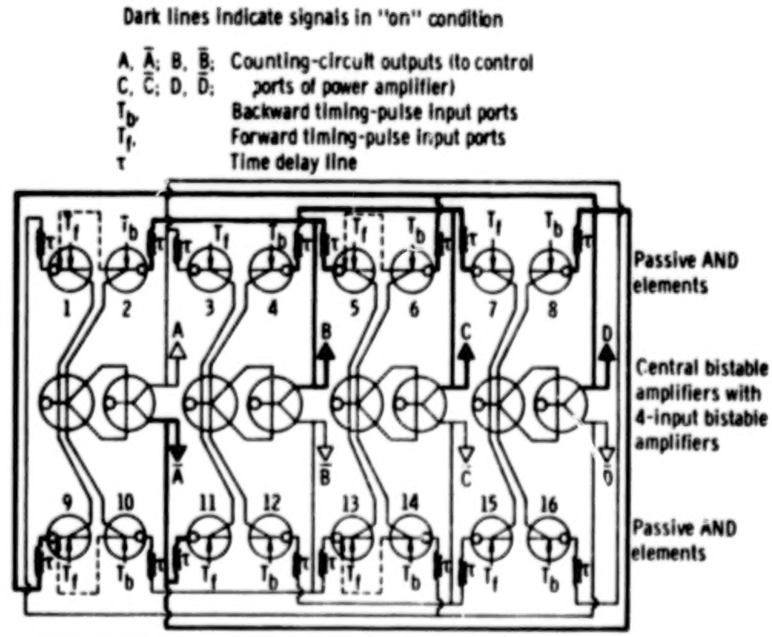
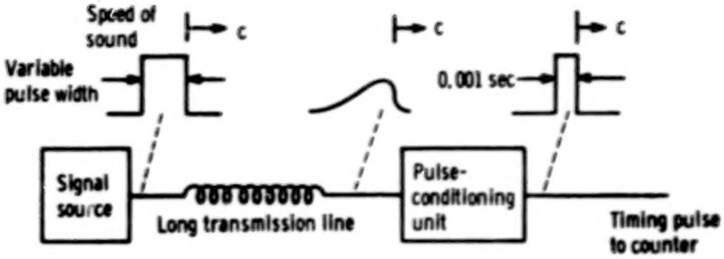
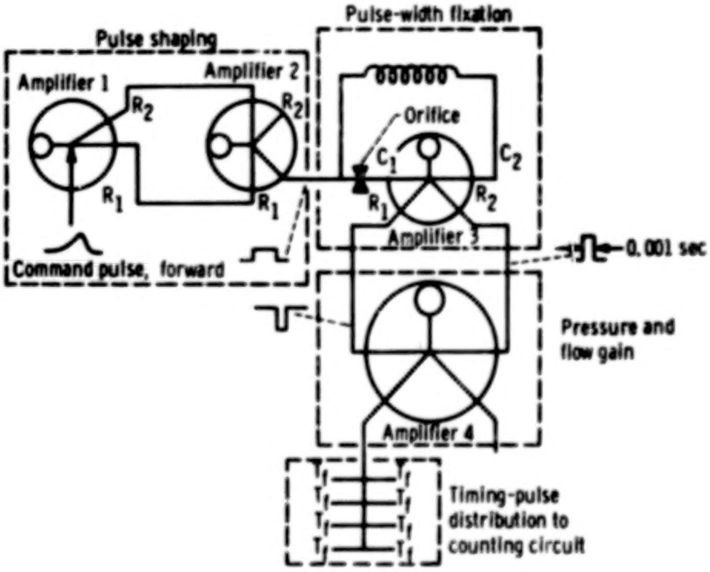

Dark lines indicate signals in "on" condition
A, Ā; B, B̄; C, C̄; D, D̄: Counting-circuit outputs (to control ports of power amplifier)
T_b, Backward timing-pulse input ports
T_f, Forward timing-pulse input ports
τ Time delay line
Passive AND elements
Central bistable amplifiers with 4-input bistable amplifiers
Passive AND elements
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
A
B
C
D
Ā
B̄
C̄
D̄
T_f
T_b
τ

Figure 3.- Counting circuit of stepping motor actuator


Speed of sound
c
c
c
Variable pulse width
0.001 sec
Signal source
Long transmission line
Pulse-conditioning unit
Timing pulse to counter

(a) Command pulse wave forms.


Pulse-width fixation
Pulse shaping
Amplifier 1
Amplifier 2
R_2
R_2
R_1
R_1
Orifice
C_1
C_2
R_1
R_2
Amplifier 3
0.001 sec
Command pulse, forward
Pressure and flow gain
Amplifier 4
T_f
Timing-pulse distribution to counting circuit

(b) Pulse-conditioning circuit.

Figure 4.- Pulse-conditioning unit

Power amplifiers are required to boost the outputs of the counting circuit to a level where they can drive the nutating gear in the stepping motor. A supersonic amplifier developed for this purpose is described in reference 3 and is shown in Figure 5.

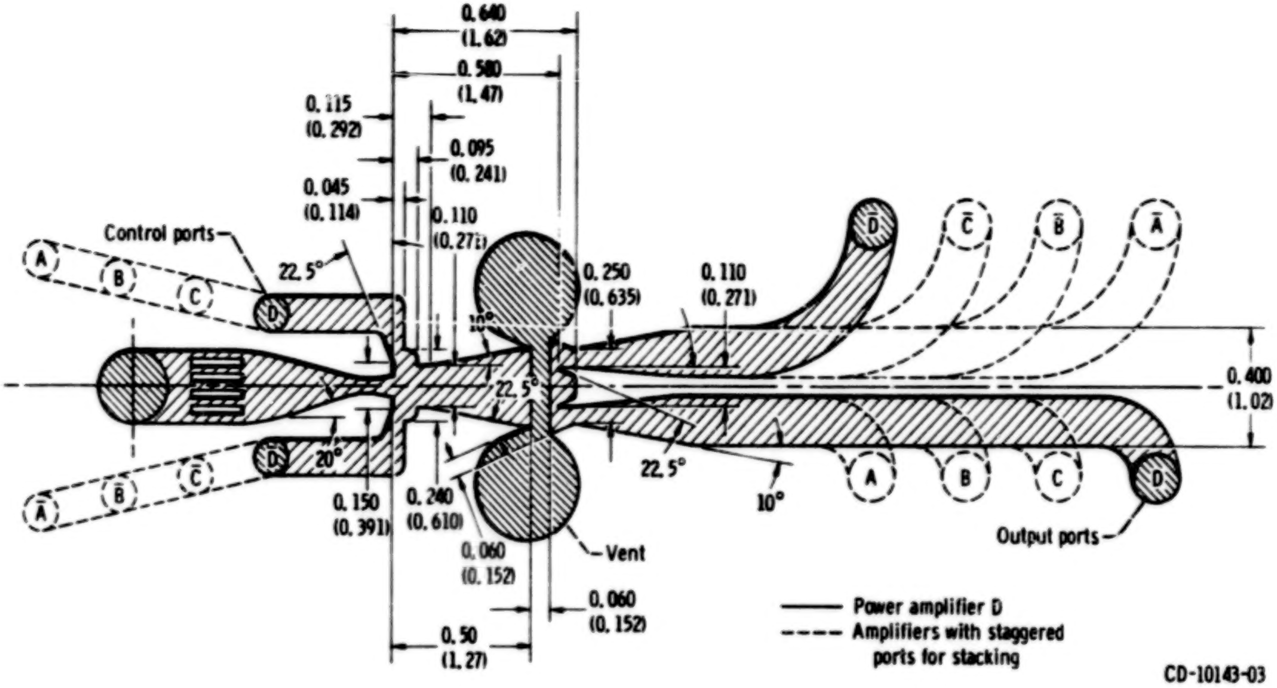
The integrated circuit consists of a stack of plates as shown in the schematic of Figure 6. One of these plates contains the two pulse-conditioning circuits; four plates contain the counting circuit; and separate plates each contain one of the four power amplifiers. Other plates in the stack contain interconnecting channels.

The plates containing the pulse conditioning circuits and the counting circuit were fabricated from a photo-etched ceramic. The interconnecting manifolds and power amplifier plates were fabricated by milling passages into acrylic plastic plates. Figure 7 shows the complete stacked assembly bolted to the bellows distribution plate. This plate then mounts on the stepping motor back plate.

The pulse-conditioning circuit plate is shown in Figure 8. Both forward and backward pulse-conditioning circuits are in the same plate. The amplifier and port identification symbols agree with those used in Figure 4. Note the long channel M which furnishes the 1-millisecond delay required to fix the command pulse width.

The performance of the integrated circuit equalled or surpassed that of the breadboard circuit. The maximum speed was 300 steps per second and maximum torque 90 in.-lb (1000 cm-N)

If the stepping motor is to be used in an analog control system, a conversion from an input pressure to pulses, the frequency of which is proportional to that input signal, is required. Reference 4 describes an analog-to-pulse-frequency converter which is intended to fulfill this function. The characteristics of the converter must be such that the frequency of the output pulses is proportional to the deviation of the input pressure from a null value as shown in Figure 9. When the input pressure is increased above the null value, the output pulse frequency increases and appears at the forward pulse output. As the input pressure decreases from the null value, the output frequency also increases but appears at the backward pulse output. The circuit saturates at a maximum frequency in either the forward or backward direction and remains at that maximum frequency for input pressures greater or less than their normal range.


0.640
(1.62)
0.580
(1.47)
0.115
(0.292)
0.095
(0.241)
0.045
(0.114)
0.110
(0.271)
Control ports
22.5°
10°
22.5°
0.250
(0.635)
0.110
(0.271)
A
B
C
D
Ā
B̄
C̄
D̄
0.400
(1.02)
20°
0.150
(0.391)
0.240
(0.610)
0.060
(0.152)
Vent
22.5°
10°
Output ports
0.060
(0.152)
0.50
(1.27)
Power amplifier D
Amplifiers with staggered ports for stacking
CD-10143-03

(a) Engraving pattern. Depth of engraving (except for vents), 0.060 inch (0.153 cm); Mach 1.63 nozzle (throat, 0.04 in. (0.102 cm). All dimensions are in inches (cm).

Figure 5.- Power amplifiers used in integrated circuits A and B

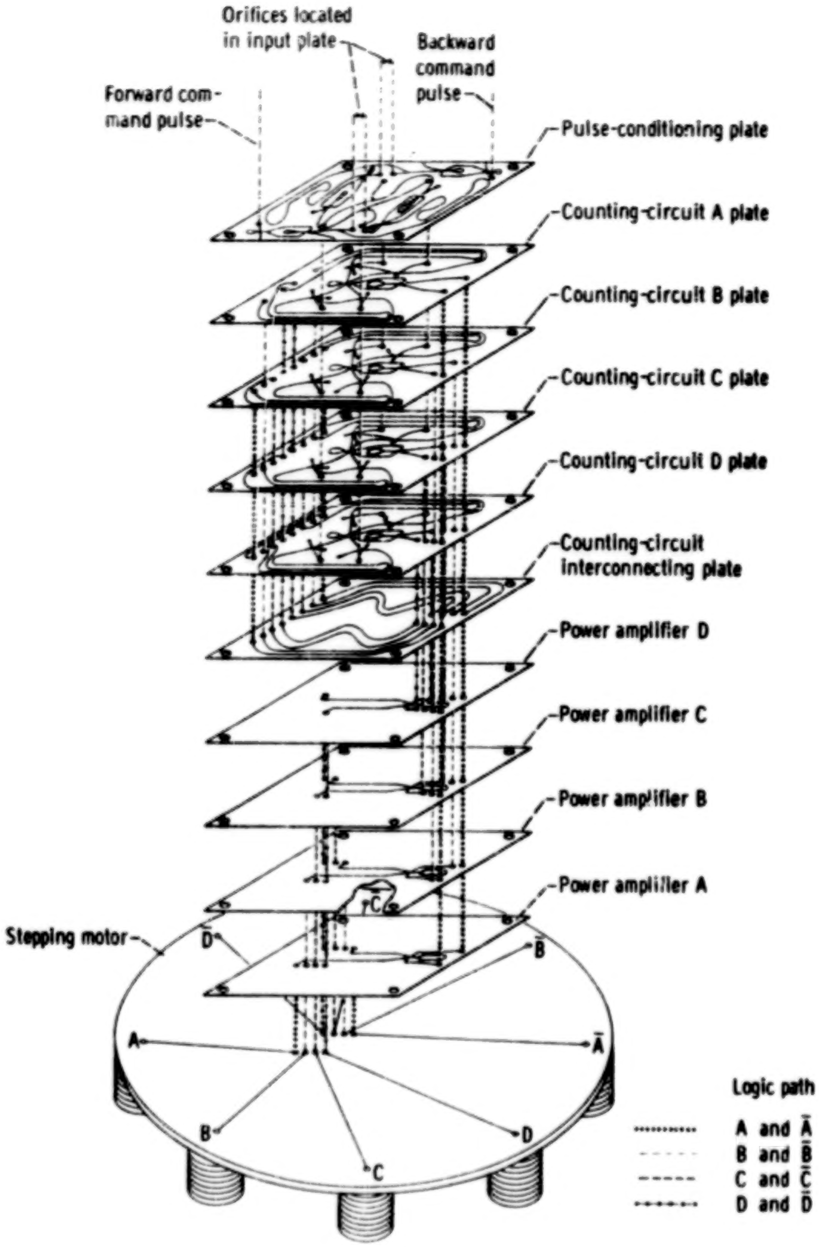

Orifices located in input plate
Backward command pulse
Forward command pulse
Pulse-conditioning plate
Counting-circuit A plate
Counting-circuit B plate
Counting-circuit C plate
Counting-circuit D plate
Counting-circuit interconnecting plate
Power amplifier D
Power amplifier C
Power amplifier B
Power amplifier A
Stepping motor
C
$\bar{D}$
$\bar{B}$
A
$\bar{A}$
B
C
D
Logic path
A and $\bar{A}$
B and $\bar{B}$
C and $\bar{C}$
D and $\bar{D}$

Figure 6.- Expanded schematic diagram of integrated stepping motor circuit

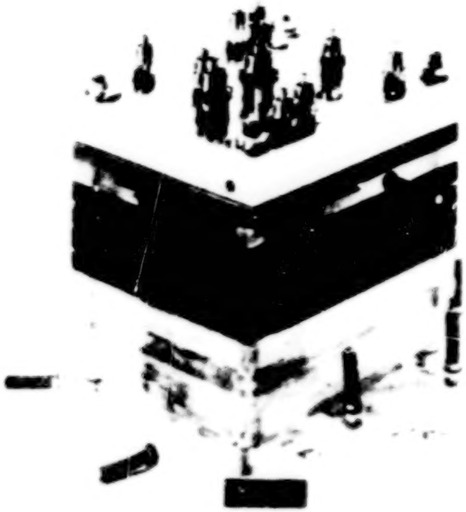


Figure 7.- Integrated circuit mounted on power amplifiers and bellows distribution plate

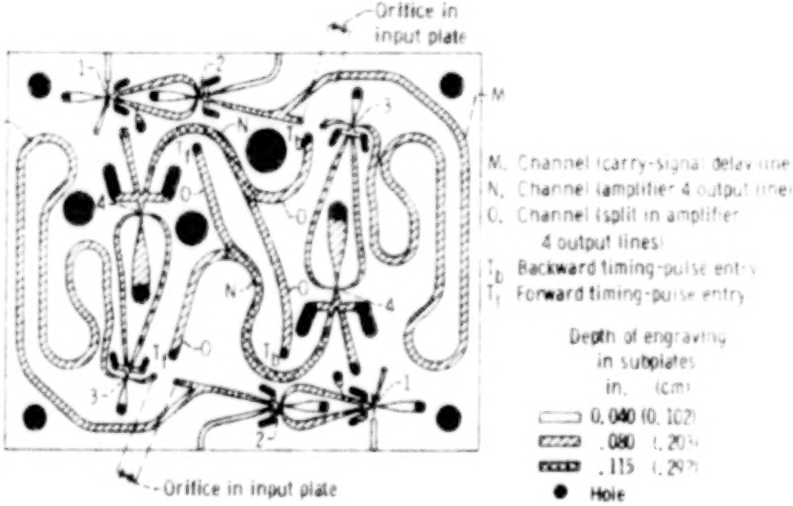

Orifice in input plate
1
2
3
4
M
N
O
T_b
T_f
Orifice in input plate
M, Channel (carry-signal delay line)
N, Channel (amplifier 4 output line)
O, Channel (split in amplifier 4 output lines)
T_b Backward timing-pulse entry
T_f Forward timing-pulse entry
Depth of engraving in subplates
in. (cm)
0.040 (0.102)
.080 (.203)
.115 (.292)
Hole

Figure 8.- Pulse conditioning plate

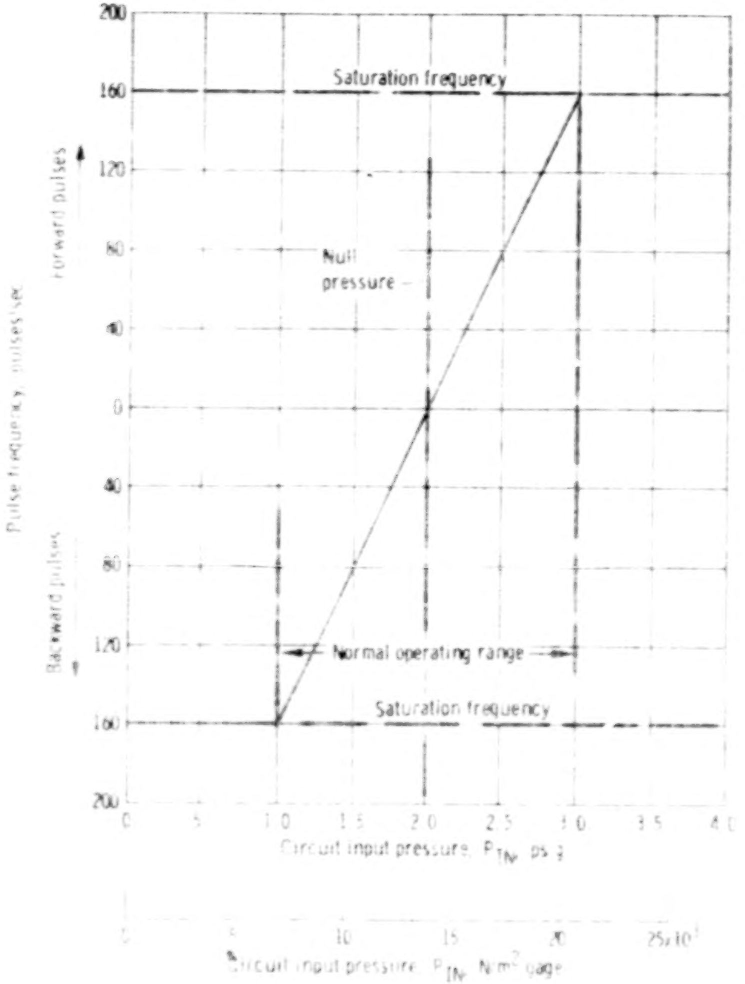


Figure 9.- Desired analog-to-pulse-frequency converter steady-state characteristics

The analog-to-pulse frequency converter is made up of three subcircuits as shown in Figure 10:

1. Oscillator circuit

2. Conditioning circuit

3. Output selector circuit

Oscillator Circuit

The oscillator circuit is composed of three bistable amplifiers with R-C time lags in the feedback paths. Amplifier A-1 outputs are connected through resistances and volumes to the control ports of two other amplifiers, A-2 and A-3. One output of each amplifier A-2 and A-3 is then connected to the control ports of amplifier A-1. The input pressure P_2 to the oscillator circuit is connected to both the remaining control ports of amplifiers A-2 and A-3.

To follow the sequence of operation of the oscillator circuit, assume that the output of amplifier A-1 is switched to resistance R_2. The pressure in volume C_2 rises in a manner approximating a first-order time constant system. When the pressure in volume C_2 is slightly higher than the oscillator input pressure P_2, amplifier A-2 will switch to the control port of amplifier A-1. This will switch amplifier A-1 to resistance R_3. As the pressure in volume C_2 drops below the oscillator input pressure P_2, amplifier A-2 will switch from the control port of amplifier A-1. However, since amplifier A-1 is bistable, its output will remain switched to R_3. Similarly, the pressure in volume C_3 rises until it is slightly higher than the oscillator input pressure P_2, at which time amplifier A-2 switches to the control port of amplifier A-1. Amplifier A-1 then switches to R_2. When the pressure in C_3 drops below the oscillator input pressure P_2, amplifier A-3 switches from the control port of amplifier A-1. This completes one cycle.

As the input pressure P_2 increases, more time is required for the volume pressures to exceed P_2. This increased time decreases the pulse frequency. The frequency becomes zero when the input pressure P_2 equals the maximum obtainable volume pressure.

Conditioning Circuit

The conditioning circuit is required to provide an input to the oscillator circuit that is a maximum when P_{IN} is at the null

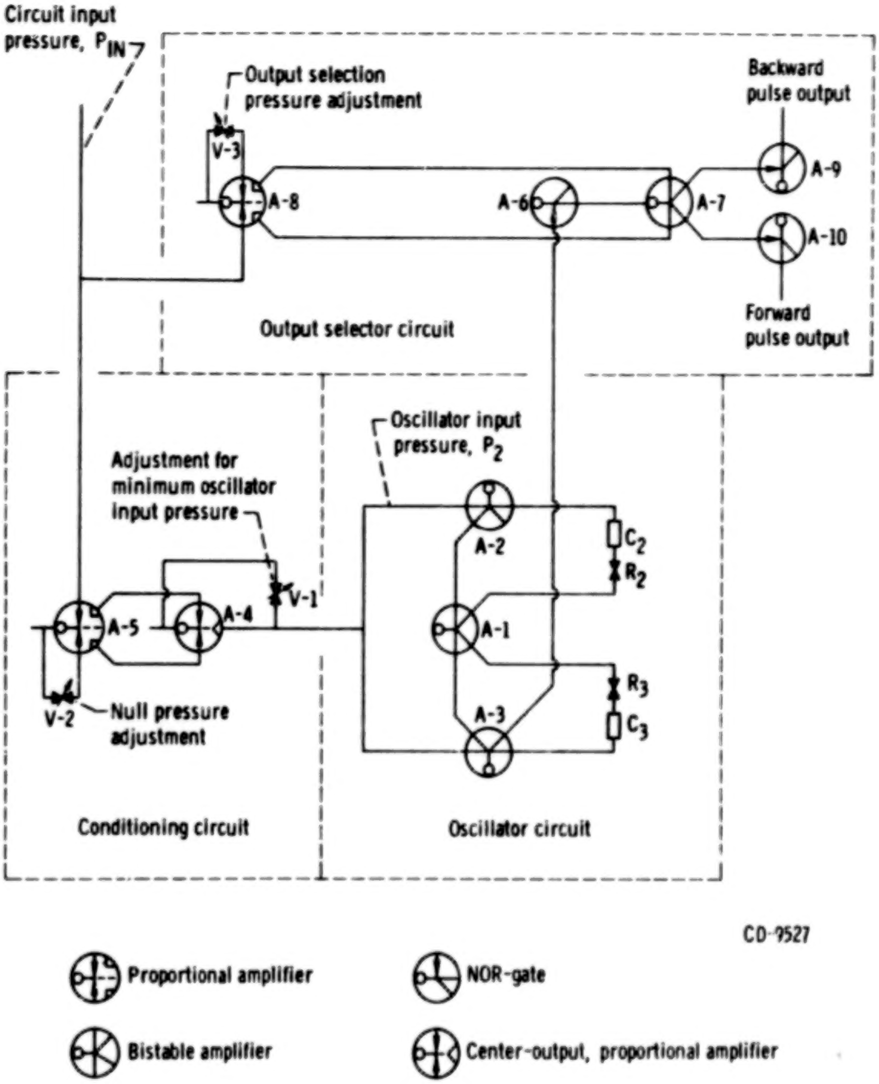

Circuit input pressure, P_IN
Output selection pressure adjustment
V-3
A-8
A-6
A-7
Backward pulse output
A-9
A-10
Forward pulse output
Output selector circuit
Oscillator input pressure, P2
Adjustment for minimum oscillator input pressure
A-2
C2
R2
V-1
A-5
A-4
A-1
R3
C3
A-3
V-2
Null pressure adjustment
Conditioning circuit
Oscillator circuit
CD-9527
Proportional amplifier
NOR-gate
Bistable amplifier
Center-output, proportional amplifier

Figure 10.- Schematic diagram of analog-to-pulse-frequency converter

value and decreases when the input pressure P_{IN} to the converter either increases or decreases from the null value. The maximum conditioning-circuit output pressure corresponds to that oscillator circuit input pressure P_2 which causes the pulse frequency to be zero. The minimum output pressure from the conditioning circuit corresponds to the oscillator circuit input pressure, which causes the pulse frequency of the oscillator circuit to be the desired maximum value. To provide flat saturation characteristics, the minimum conditioning-circuit output pressure must not fall below this value.

The conditioning circuit consists of two proportional amplifiers. Amplifier A-4 is a center-output amplifier that is not available commercially. An additional flow source is provided through valve V-1. This source maintains the output pressure of amplifier A-4 at the desired minimum pressure. Valve V-2 is used to set the null pressure at the desired value.

Output Selector Circuit

The output selector circuit directs the oscillator output pulses to either the backward output port or the forward output port. If, as shown in Figure 9, the converter input pressure P_{IN} is greater than the null pressure, the pulses appear at the forward pulse output port. If the input pressure P_{IN} is less than the null pressure, the pulses appear at the backward pulse output port.

Amplifier A-6 is a NOR-gate that serves to amplify the pulses produced by the oscillator circuit. The output of amplifier A-6 is applied to the supply port of bistable amplifier A-7. The pulses are directed toward either the forward pulse output port or the backward pulse output port, depending upon which control port pressure of amplifier A-7 is greater. Proportional amplifier A-8, which provides the control input pressures to amplifier A-7, serves to increase the switching sensitivity of the selector circuit. Valve V-3 is set so that amplifier A-7 switches when P_{IN} equals the null pressure. NOR-gates A-9 and A-10 amplify the output pulse pressures.

The output pulse frequency of the breadboard analog-to-pulse-frequency converter is plotted in Figure 11 as a function of input pressure P_{IN}.

SHOCK POSITION SENSOR

Efficient mixed-compression inlets are of interest for supersonic cruise aircraft where high recoveries are required.

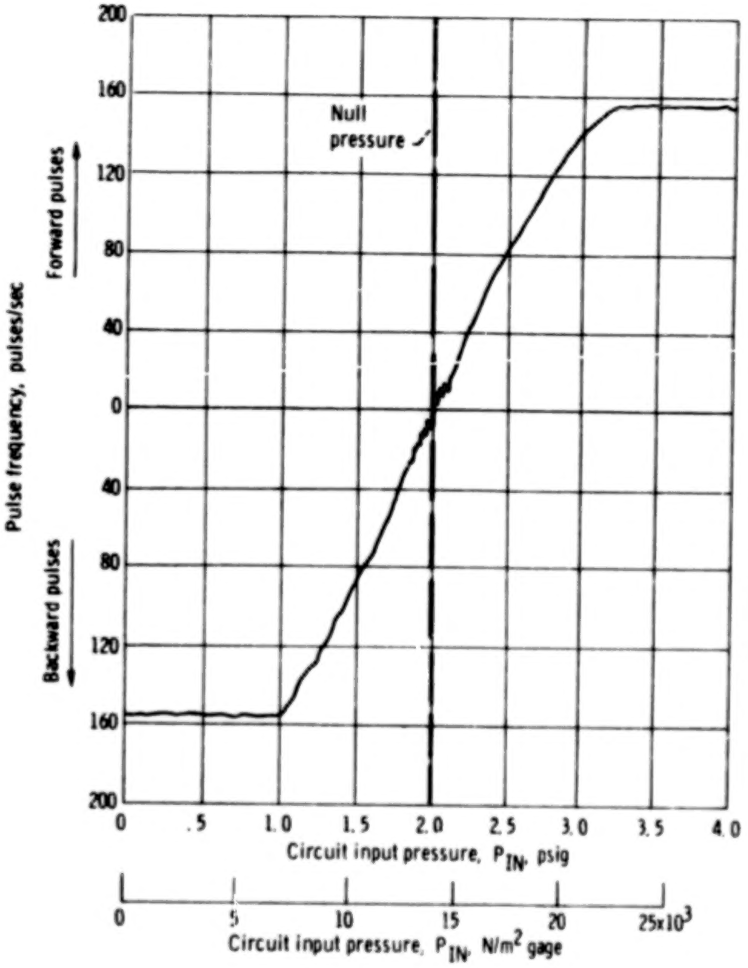


Figure 11.- Analog-to-pulse-frequency converter steady-state performance

To achieve such inlets, it is necessary to operate with the shock located close to the throat as shown in Figure 12. With the shock in this position, slight atmospheric disturbances or engine air flow changes can cause the shock to be rapidly expelled out the front of the inlet. A fast operating inlet control system is needed to prevent this from happening.

For high Mach number inlets, electronic control systems have been plagued with problems associated with sensor failure. Failures are largely due to high temperature and high vibration conditions. It is thought that a judicious choice of fluidic sensors, electronic controls, and hydraulic actuators might be ideal. A fluidic sensor can use the flow in the inlet for a power source; thus, it can be made a completely self-contained sensor located in the structure of the inlet cowl or centerbody. The speed of response is adequate for shock position control.

The fluidic shock position sensor presented here uses the wall static pressure gradient to determine shock position. Figure 13 shows wall static pressure distributions for several different shock positions. These data were taken from a Mach-2.5 supersonic inlet which was tested in the 10- by 10-foot supersonic wind tunnel at the Lewis Research Center. From the figure we can see that the position of the shock can be estimated by eye. The shock is preceded by a neutral or slightly negative slope and followed by a steep positive slope. Thus, for the pressure distribution represented by the symbol D in Figure 13 the shock would be between taps B and C. For the pressure distribution represented by the symbol ◇ , the shock would be between taps D and E.

The sensor's logic diagram is shown in Figure 14. The sensor uses six pressure taps in the inlet just downstream of the throat. The first state of the diagram is a row of comparators. The shock position is determined by comparing three adjacent pressures to determine if a minimum exists. Each comparator has two outputs. The output turned "on" depends upon which of the two inputs is largest. The second row is composed of logical AND elements. Their outputs are on only when both of their inputs are on, which is true when there is a minimum for that element.

The sensor is built so that all outputs upstream of the shock are off and outputs down stream of the shock are on. This is done so that a stepwise proportional output signal can be obtained by simply adding the output signals. This is done by arranging the last row, the output elements, so that when any one element has an input signal, its output is fed to the next element upstream and turns it on. The heavy lines in Figure 14 indicate the signals that are activated when the shock is between C and D.

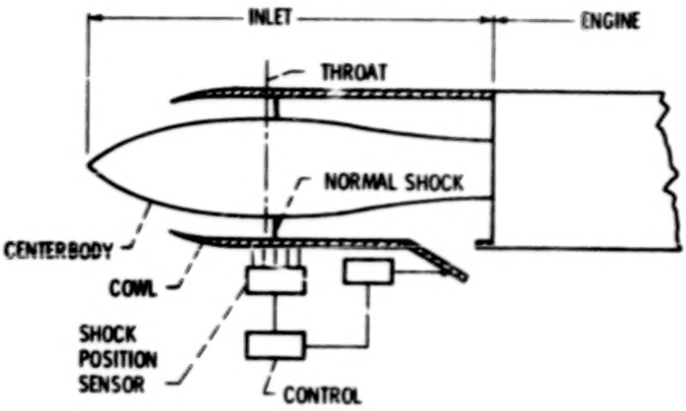

INLET
ENGINE
THROAT
NORMAL SHOCK
CENTERBODY
COWL
SHOCK POSITION SENSOR
CONTROL

Figure 12.- Schematic of supersonic jet engine inlet

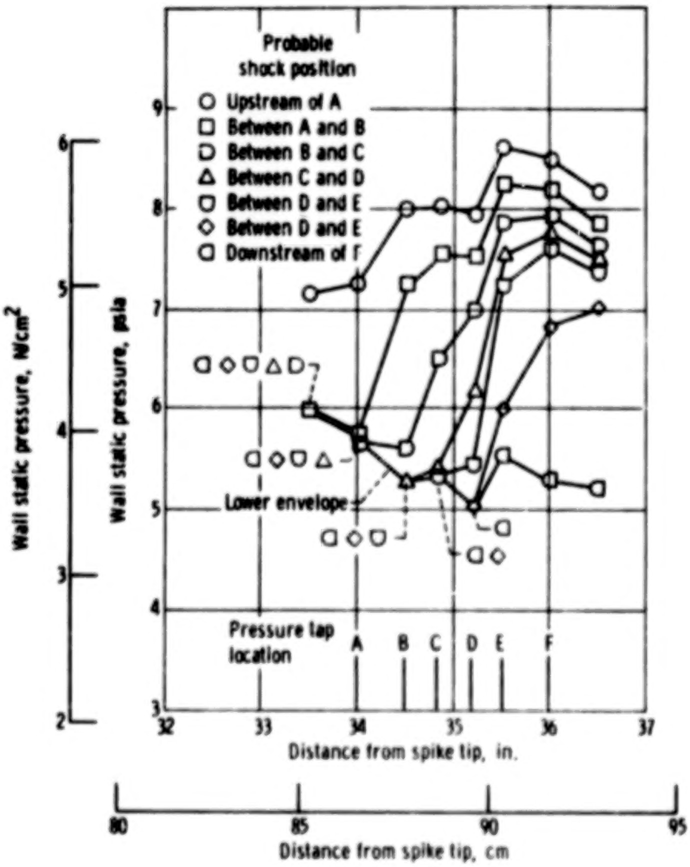


Figure 13.- Typical wall static pressure distributions as function of shock position

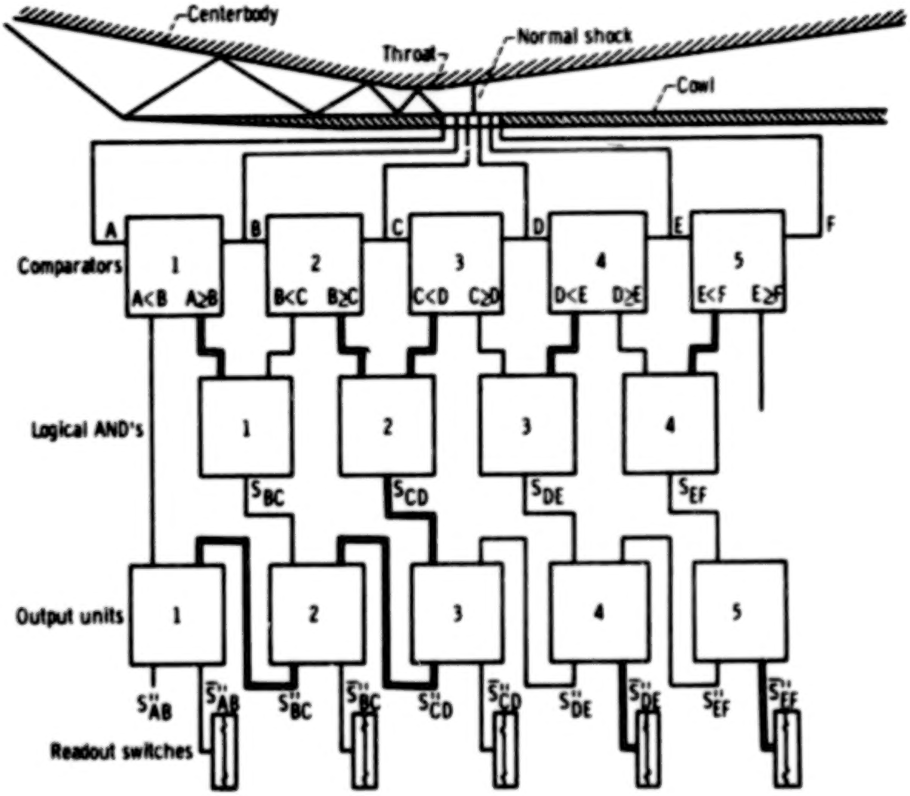

Centerbody
Throat
Normal shock
Cowl
A
B
C
D
E
F
Comparators
1
2
3
4
5
A<B
A≥B
B<C
B≥C
C<D
C≥D
D<E
D≥E
E<F
E≥F
Logical AND's
1
2
3
4
S_{BC}
S_{CD}
S_{DE}
S_{EF}
Output units
1
2
3
4
5
S''_{AB}
$\bar{S}''_{AB}$
S''_{BC}
$\bar{S}''_{BC}$
S''_{CD}
$\bar{S}''_{CD}$
S''_{DE}
$\bar{S}''_{DE}$
S''_{EF}
$\bar{S}''_{EF}$
Readout switches

Figure 14.- Block diagram of shock position sensor

A diagram of the fluidic sensor is shown in Figure 15. The comparator units are made of OR-NOR elements in which the vents are used as one of the two inputs. The next row is made up of OR-NOR elements used as passive AND elements. There is an output only if both inputs are on. The output units are also OR-NOR elements, but are used in the usual manner.

The fluidic sensor was run on two occasions in the Mach-2.5 supersonic inlet During the first tests the sensor performed reasonably well until the unit got contaminated by material from the tunnel driers. A filter was added to the supply line and the tests were repeated. This time the sensor correctly indicated the profile minimums, but in some cases it read falsely because of false minimums.

Another fluidic sensor is being built which indicates the shock position from the first positive slope of the pressure profile. The false minimums will be biased out by bleeding supply air into the sensor input lines.

VORTEX VALVE SHOCK STABILIZATION

Another program related to controlling the position of the normal shock with fluidic devices is a vortex valve shock stabilization program. This system uses a series of vortex valves positioned around the inlet.

Figure 16 (a) shows a cutaway view of a typical vortex valve. The valve consists of a cylindrical cavity in which flow enters both radially and tangentially. Flow exists from the hole in the center. If flow enters only through the radial port, it goes directly to the exit hole. Then, if tangential flow is added, a swirl is introduced to the fluid. The resulting vortex produces a resistance to flow, reducing the flow from the outlet. Figure 16(b) shows the total flow through the valve plotted as a function of pressure ratio. If there is a small radial pressure with respect to tangential pressure, a strong vortex is created with little total flow. As the radial pressure is increased, there is no increase in total flow until the pressure ratio is about 0.6. Then there is a sharp increase in flow as radial pressure is increased still further.

Figure 17 displays a schematic of the vortex valve stabilization system installed in an inlet. The radial port is connected to the inlet just downstream of the throat. The tangential port is connected to the inlet well downstream of the shock.

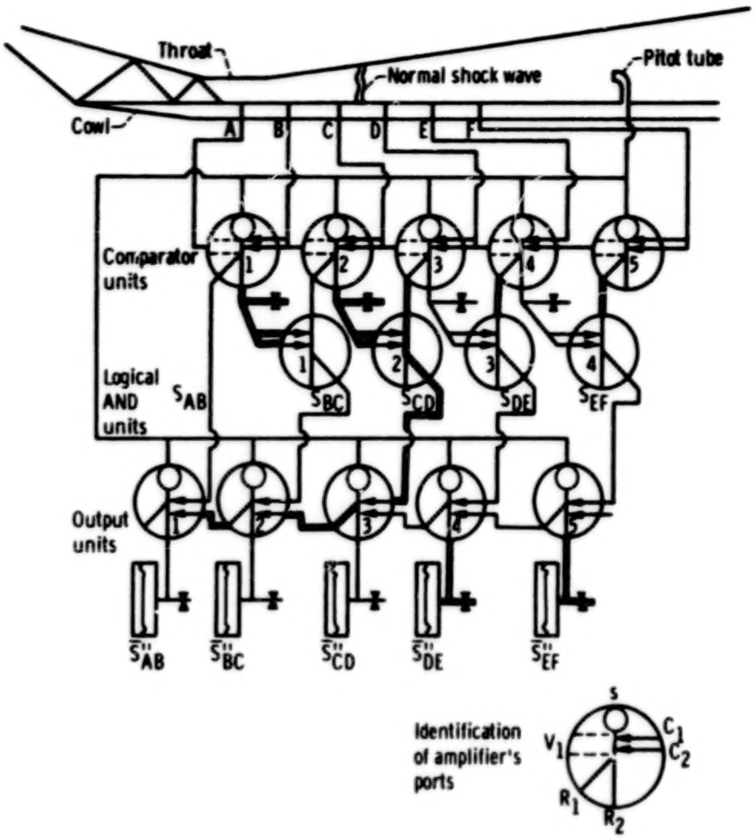
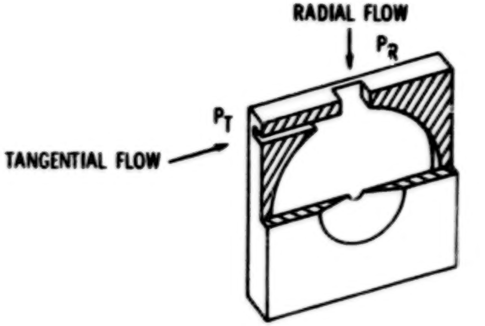
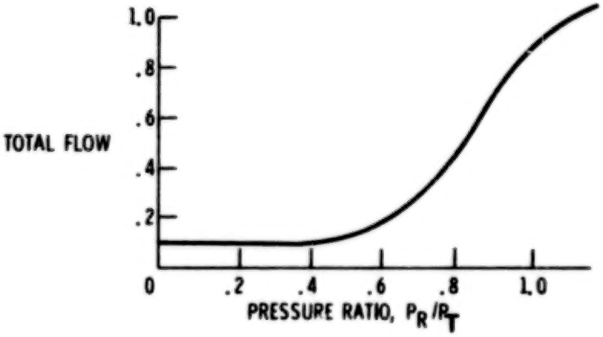

Throat
Normal shock wave
Pitot tube
Cowl
A
B
C
D
E
F
Comparator units
1
2
3
4
5
Logical AND units
S_{AB}
1
2
3
4
S_{BC}
S_{CD}
S_{DE}
S_{EF}
Output units
1
2
3
4
5
$\bar{S}''_{AB}$
$\bar{S}''_{BC}$
$\bar{S}''_{CD}$
$\bar{S}''_{DE}$
$\bar{S}''_{EF}$
Identification of amplifier's ports
S
C_1
C_2
V_1
R_1
R_2

Figure 15.- Schematic of fluidic shock position sensor


RADIAL FLOW
P_R
P_T
TANGENTIAL FLOW

(A) VORTEX VALVE.



(B) FLOW AS A FUNCTION OF PRESSURE RATIO FOR A TYPICAL VORTEX VALVE.

Figure 16.- Typical vortex value and its pressure-flow characteristics

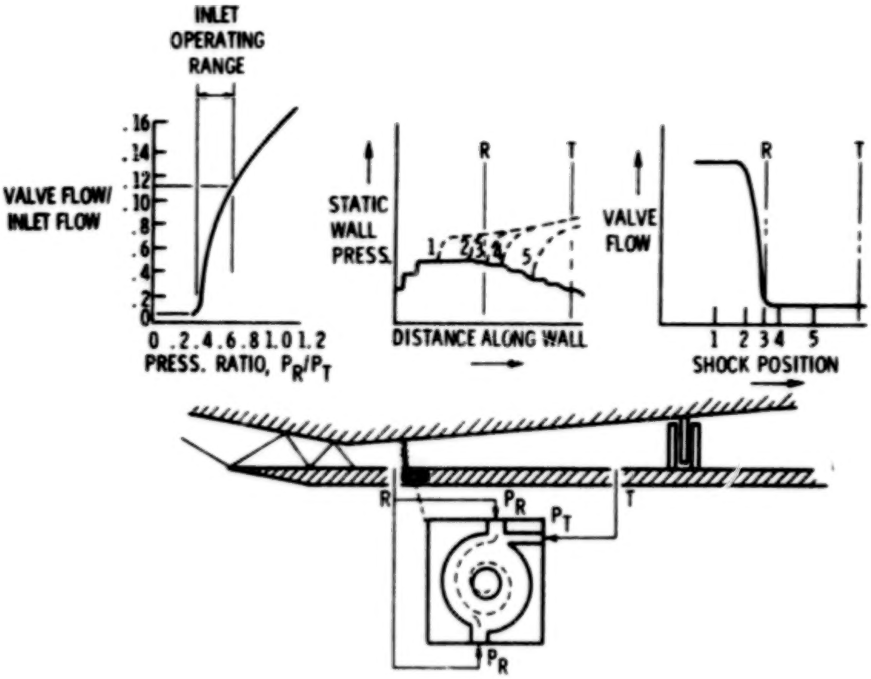

INLET OPERATING RANGE
VALVE FLOW/ INLET FLOW
.16
.14
.12
.10
.8
.6
.4
.2
0
0 .2 .4 .6 .8 1.0 1.2
PRESS. RATIO, P_R/P_T
STATIC WALL PRESS.
R
T
1
2,3,4
5
DISTANCE ALONG WALL
VALVE FLOW
R
T
1 2 3 4 5
SHOCK POSITION
R
P_R
P_T
T
P_R

Figure 17.- Shock stabilization system for supersonic jet engine inlet

The center graph shows typical pressure profiles for various shock locations. The left graph shows the vortex valve flow as a function of pressure ratio. It should be noted that the characteristics are different from those of the typical vortex valve shown in Figure 16. Since the ratio of radial pressure to tangential pressure encountered in the inlet varies from about 0.4 to 0.6, it was necessary to alter the vortex valve characteristics so that the high flow gain region occurs over the same pressure ratio range. The method used to alter the valve characteristics to match the inlet characteristics is discussed later.

When the shock is in a stable location downstream of the radial tap, the pressure ratio is low, the valve is vortexing and flow through the valve is low. As the shock moves upstream and passes over the radial tap, there is an increase in pressure ratio, resulting in a weaker vortex and, therefore, more flow. If the shock wave moves further upstream, the flow increases still further, thus bypassing more flow around the engine, keeping the shock in a stable region. The right hand graph shows the increase in flow as the shock moves upstream.

The modifications required to alter the vortex valve operating characteristics are shown in Figure 18. In the typical vortex valve configuration discussed earlier, the radial ports are located in line with the center output port. With no tangential flow the radial flow is directly out the output port. In the modified vortex valve the radial ports are offset with respect to the outlet port. Now, with no tangential flow a vortex is set up in the opposite direction to the normal vortex produced by the tangential port. In effect, a bias is created which moves the operating region to lower pressure ratio values. It also increases the gain of the element.

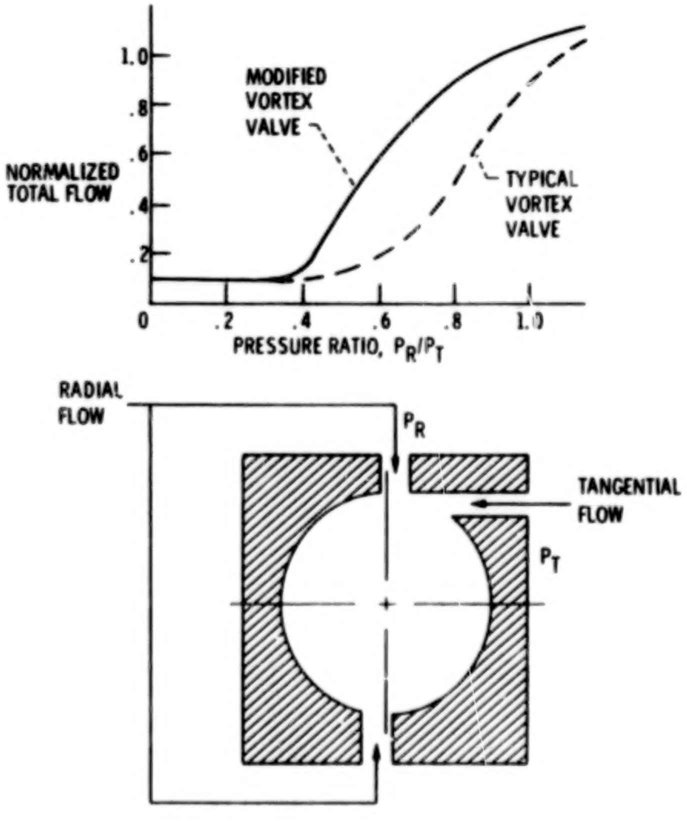

NORMALIZED TOTAL FLOW
1.0
.8
.6
.4
.2
0
.2
.4
.6
.8
1.0
PRESSURE RATIO, P_R/P_T
MODIFIED VORTEX VALVE
TYPICAL VORTEX VALVE
RADIAL FLOW
P_R
TANGENTIAL FLOW
P_T

Figure 18.- Modified vortex value that matches inlet characteristics

REFERENCES

1. Howland, G. R.: "Pneumatic Nutator Actuator Motor." Rep. No. BPAD-863-16719R, NASA CR-54788, Bendix Corp., Oct. 17, 1965.

2. Griffin, William S.: "A Breadboard Flueric-Controlled Pneumatic Stepping Motor System." NASA TN D-4495, 1968.

3. Dustin, M. O.; and Wallhagen, R. E.: "Design and Performance of Two Integrated Circuits for Fluidic-Controlled Pneumatic Stepping Motor System." NASA TN D-5155, 1969.

4. Dustin, M.O.: "Design and Performance of Flueric Analog-to-Pulse-Frequency Converter." NASA TN D-4497, 1968.

5. Griffin, W. S.: "Design and Performance of a Flueric Shock Position Sensor for a Mixed-Compression Supersonic Inlet." NASA TM X-1733, 1969.

MINIMUM ENERGY CONTROL OF A CLASS OF ELECTRICALLY DRIVEN VEHICLES

By Yilmaz E. Sahinkaya
Jet Propulsion Laboratory

SUMMARY

This paper presents theoretical and experimental results obtained in the investigation of a minimum energy control problem associated with a class of electrically driven vehicles. Analytical results are obtained by making several justifiable approximations in the dynamical equations of the plant, the performance index of which is related to the minimization of system energy consumption for any required control action.

The control problem of interest is simplified for practical reasons and solved by using Bellman's invariant imbedding technique. The optimality of the resulting control law is compared with those of the bang-bang control law, which is the exact solution for the case of speed-setting control action, and the classical control laws under identical conditions. A stochastic optimization problem is then formulated and solved by using Bellman's dynamic programming technique.

The resulting control law is mechanized for a particular vehicle configuration which is simulated in the laboratory. Silicon-controlled rectifiers are chosen over power transistors to control the motor armature current to determine a flexible minimum energy controller configuration which can be used for vehicles ranging from fractional to integral horse power ratings. A transistorized pulse-width modulator is utilized to control silicon-controlled rectifiers according to the control law. The theoretical and practical results clearly display a significant effort which is exercised in bridging the gap between theory and practice.

INTRODUCTION

In recent years, following the pioneering works of Pontryagin, Bellman, and Kalman, an intense amount of research has been carried out by many workers in the area of modern control theory. The optimization techniques of modern control theory, along with digital computers, provide a new approach for the design of automatic control systems.

One of the most challenging tasks in the formulation of a practical problem in the context of optimal control theory is the selection of an appropriate performance index. Even though an adequate mathematical description of the plant is available, the quality of the results obtained by the application of optimal control theory is dependent upon the selection of a meaningful performance index.

This investigation treats the formulation and solution of minimum energy control problems associated with a class of electrically driven vehicles in the context of modern control theory and the mechanization of the resulting control law in the light of current engineering practice. The essential elements of the minimum energy control system resulting from the implementation of the control law are illustrated in Figure 1.

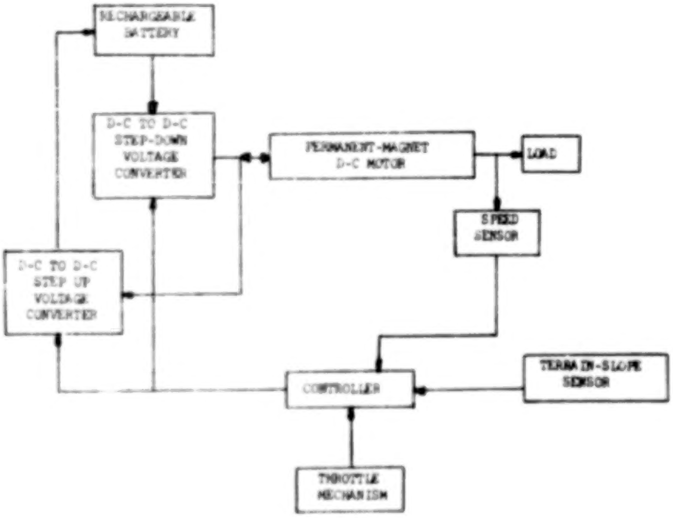

RECHARGEABLE BATTERY
D-C TO D-C STEP-DOWN VOLTAGE CONVERTER
PERMANENT-MAGNET D-C MOTOR
LOAD
SPEED SENSOR
D-C TO D-C STEP UP VOLTAGE CONVERTER
CONTROLLER
TERRAIN-SLOPE SENSOR
THROTTLE MECHANISM

Figure 1.- Essential elements of the minimum energy control system

Developing a reasonably accurate mathematical model for this specific control system is a necessity before the theoretical concepts of optimal control theory can be applied to synthesize an optimal controller. The following procedure has been used in the determination of the equations of the system to be controlled, i.e., plant.

1. Obtain the expressions for the total kinetic, potential, and dissipation energies of the system which consists of several electrical and mechanical interacting components;

2. Select the suitable generalized coordinates for the expressions obtained in Eq. (1) and use the auxiliary expressions for the generalized forces to write down the system equations according to Lagrange's energy method.

The plant equations are determined to be:

$$\dot{x}_1(t) = -\left(\frac{f}{J}\right)x_1(t) + \left(\frac{k_t}{J}\right)x_2(t) - \left(\frac{1}{J}\right)v(t) \tag{1}$$

$$\dot{x}_2(t) = -\left(\frac{k_b}{\ell}\right)x_1(t) - \left(\frac{r}{\ell}\right)x_2(t) + \left(\frac{1}{\ell}\right)u(t) \tag{2}$$

where

$\dot{x}_1(t)$ $\triangleq$ angular acceleration of the electric motor [rad/sec^2]

$x_1(t)$ $\triangleq$ angular velocity of the electric motor [rad/sec]

$\dot{x}_2(t)$ $\triangleq$ derivative of motor current [amps/sec]

$x_2(t)$ $\triangleq$ motor current [amps]

$v(t)$ $\triangleq$ total disturbance torque as referred to motor [newton-meter]

$u(t)$ $\triangleq$ control voltage [volts]

J $\triangleq$ total system inertia as referred to motor [newton-meter/rad/sec^2]

f $\triangleq$ total system damping coefficient as referred to motor [newton-meter/rad/sec]

k_t $\triangleq$ motor torque constant [newton-meter/amps]

k_b $\triangleq$ motor back emf constant [volts/rad/sec]

ℓ $\triangleq$ motor inductance [henries]

r $\triangleq$ motor resistance [ohms].

In the derivation of the set of differential equations denoted by Eqs. (1) and (2) the following assumptions have been made:

1. The electric motor propulsion force is transmitted to vehicle's wheel through a speed-reducer and no slipping takes place between the tires and the surface of terrain;

2. The aerodynamic drag force is negligible, i.e., low-speed operation;

3. The mechanical system parameters, J and f, and the electrical system parameters, ℓ and r, are constants;

4. The motor operation is unsaturated.

The system electrical energy consumption is selected to be the design criterion or performance index for this optimization problem. The system performance index is denoted by:

$$E = \int_{0}^{T} u(t)\ x_2(t)\ dt \tag{3}$$

where

$T \overset{\Delta}{=}$ response time.

The integrand in Eq. (3) represents the electric power flow which can flow from a rechargeable battery system into the motor circuit or from the motor circuit into the electric energy source over different intervals of time during the controlling process. The set of boundary conditions to be satisfied by the state variables $x_1(t)$ and $x_2(t)$ for the three cases of control action are given below:

Case 1 - Speed-control.

$$x_1(0) = x_{01} = \alpha_1 \quad , \quad x_1(T) = \alpha_1$$

$$x_2(0) = x_{02} \quad , \quad x_2(T) = \alpha_2$$

$$\alpha_2 \text{ is chosen such that } \dot{x}_1(T) = 0 \tag{4}$$

$$v(0-) \neq v(0) \quad , \quad v(T) = \beta$$

$$v(t) = \beta \quad , \quad 0 \leq t \leq T$$

Case 2 - Speed-setting.

$$x_1(0) = x_{01} \neq \alpha_1 \quad , \quad x_1(T) = \alpha_1$$

$$x_2(0) = x_{02} \quad , \quad x_2(T) = \alpha_2$$

$$\alpha_2 \text{ is chosen such that } \dot{x}_1(T) = 0 \tag{5}$$

$$v(0-) = v(0) \quad , \quad v(T) = \beta$$

$$v(t) = \beta \quad , \quad 0 \leq t \leq T$$

Case 3 - Speed-control and speed-setting.

$$x_1(0) = x_{01} \neq \alpha_1 \quad , \quad x_1(T) = \alpha_1$$

$$x_2(0) = x_{02} \quad , \quad x_2(T) = \alpha_2$$

$$\alpha_2 \text{ is chosen such that } \dot{x}_1(T) = 0 \tag{6}$$

$$v(0-) \neq v(0) \quad , \quad v(T) = \beta$$

$$v(t) = \beta \quad , \quad 0 \leq t \leq T$$

In Eqs. (4), (5), and (6), the conditions placed on $v(t)$ indicate that the disturbance torque $v(t)$ remains constant in $0 \leq t \leq T$. This means the vehicle is assumed to be traversing over a portion of terrain, which has a constant slope angle with respect to local horizontal, during the controlling interval $(0,T)$.

The control voltage $u(t)$ is constrained by the inequality

$$0 \leq u(t) \leq U \tag{7}$$

where

U = battery voltage which is assumed to be constant in $0 \leq t \leq T$.

The statement of the control problem is as follows: Given the linear time-invariant system, i.e., Eqs. (1) and (2), the performance index, i.e., Eq. (3), an arbitrary response time T, an equality constraint on u(t), i.e., Eq. (7), determine u(t) which satisfies the boundary conditions on $x_1(t)$ and $x_2(t)$ as described by Eq. (4) for speed-control, Eq. (5), for speed-setting, Eq. (6), for speed-control and speed-setting and minimizes the performance index. It is important to note that for $t \geq T$, it is required to maintain the vehicle speed constant at its terminal value, i.e., $x_1(T) = \alpha_1$, until a new disturbance or a new desired speed-setting is applied to the system.

The exact solution of the basic optimization problem is obtained by using Pontryagin's Maximum Principle (ref. 1) together with necessary conditions. As shown in reference 2, the solution yields a dual-mode control law which may consist of bang-bang and singular control actions during the controlling interval. The engineering realization of such a controller is very complex and expensive. To circumvent this difficulty the following assumptions are made on the basic optimization problem:

1. The armature inductance is neglected.
2. The inequality constraint on U(t) as given by Eq. (7) is removed.

These assumptions convert the basic optimization problem to an approximate optimization problem with a linear plant equation and a quadratic performance index as given below:

Plant:

$$\dot{x}(t) = -\left(\frac{f}{J} + \frac{k_t k_b}{Jr}\right)x(t) + \left(\frac{k_t}{Jr}\right)u(t) - \left(\frac{1}{J}\right)v(t) \tag{8}$$

where

$$x(t) \triangleq x_1(t)$$

Performance Index:

$$E = \int_0^T \frac{1}{r}\left[u^2(t) - k_b u(t)\, x\ (t)\right]dt \tag{9}$$

Boundary Conditions:

Case 1 - Speed-control.

$$\begin{array}{llll} x(0) = \alpha & , & x(t) = \alpha & \\ v(0-) \neq v(0) & , & v(t) = \beta & \quad (10) \\ v(t) = \beta & , & 0 \leq t \leq T & \end{array}$$

Case 2 - Speed-setting.

$$\begin{array}{llll} x(0) = x_o & , & x(t) = \alpha & \\ v(0-) = v(0) & , & v(t) = \beta & \quad (11) \\ v(t) = \beta & , & 0 \leq t \leq T & \end{array}$$

Case 3 - Speed-control and Speed-setting.

$$\begin{array}{lll} x(0) = x_o & , & x(T) = \alpha \\ v(0-) \neq v(0) & , & v(T) = \beta \\ v(t) = \beta & , & 0 \leq t \leq T \end{array}$$

From the observation of Eq. (9), it is easy to see that the performance index has the same fundamental characteristic as that of Eq. (3), i.e., the integrand of each integral represents the electric power flow which is reversible.

The solution of the resulting approximate optimization problem is obtained as follows:

Define the pre-Hamiltonian H:

$$H = (\frac{1}{r})u^2(t) - (\frac{k_b}{r})u(t)\ x\ (t) + \{-(\frac{f}{J} + \frac{k_t k_b}{Jr})x(t) - (\frac{1}{J})v(t) + (\frac{k_t}{Jr})u(t)\}\lambda(t) \qquad (13)$$

Application of Pontryagin's Maximum Principle yields the following results:

$$u^*(t) = (\frac{k_b}{2}) x(t) - (\frac{k_t}{2J}) \lambda(t) \quad , \quad 0 \leq t \leq T \tag{14}$$

$$H^* = -(\frac{k_b^2}{4r}) x^2(t) - (\frac{f}{J} + \frac{k_t k_b}{2Jr}) x(t)\lambda(t) - (\frac{k_t^2}{4J^2 r}) \lambda^2(t) - (\frac{1}{J}) v(t)\lambda(t) \tag{15}$$

It is seen from Eq. (14) that $u^*(t)$ is a linear function of $x(t)$ and $\lambda(t)$. Since the plant equation is linear, the performance index is quadratic, $u^*(t)$ is unique, and therefore is the optimal solution for this approximate optimization problem.

The canonic equations are then determined to be:

$$\dot{\tilde{x}}(t) = -a\,\tilde{x}(t) - b\lambda(t) - (\frac{1}{J})\beta - a\alpha \tag{16}$$

$$\dot{\lambda}(t) = c\,\tilde{x}(t) + a\lambda(t) + c\alpha \tag{17}$$

where

$$\tilde{x}(t) = x(t) - \alpha$$

$$\beta = v(t)$$

$$a = (\frac{f}{J} + \frac{k_t k_b}{2Jr}) \qquad a > 0$$

$$b = (\frac{k_t^2}{2J^2 r}) \qquad b > 0$$

$$c = (\frac{k_b^2}{2r}) \qquad c > 0$$

Since the canonic equations are linear, Bellman's invariant-imbedding technique (ref. 3) can be used to determine the feedback solution. Let $z(\eta,\tau)$ be any particular initial condition on $\tilde{x}(t)$ for a process starting at time τ with $\lambda(\tau) = \eta(\tau)$ and satisfying the boundary condition $z(\eta,T) = 0$. By referring to reference 3, it is shown that $z(\eta,\tau)$ satisfies the following first-order partial differential equation:

$$\frac{\partial z}{\partial \tau}(\eta,\tau) + \frac{\partial z}{\partial \eta}(\eta,\tau)\, g(\eta,z) = f(\eta,z) \qquad (18)$$

where

$$f(\eta,z) \triangleq -az(\tau) - b\eta(\tau) - (\frac{1}{J})\beta - a\alpha$$

$$g(\eta,z) \triangleq cz(\tau) + a\eta(\tau) + c\alpha$$

Assume a solution of Eq. (18) in the form of

$$z(\eta,\tau) = m(\tau)\,\eta(\tau) + n(\tau) \qquad (19)$$

Using Eq. (19) in Eq. (18) and collecting the coefficients of equal powers of $\eta(\tau)$ and equating them to zero yields:

$$\dot{m}(\tau) + cm^2(\tau) + 2\,am(\tau) + b = 0 \qquad (20)$$

$$m(T) = 0 \qquad (21)$$

$$\dot{n}(\tau) + (cm(\tau) + a)\, n(\tau) + \alpha cm(\tau) + a\alpha + (\frac{1}{J})\beta = 0 \qquad (22)$$

$$n(T) = 0 \qquad (23)$$

where

Eqs. (21) and (23) are obtained from Eq. (19).

Eq. (20) with its boundary condition, i.e., Eq. (21), is a first-order, ordinary, non-linear differential equation of the Riccati type, the solution of which is given by

$$m(\tau) = -b\left(\frac{\sinh\sqrt{d}\ (T-\tau)}{a \sinh\sqrt{d}\ (T-\tau) - \sqrt{d}\cosh\sqrt{d}\ (T-\tau)}\right) \qquad (24)$$

where

$$d = a^2 - bc$$

Eq. (22), with its boundary condition, i.e., Eq. (23), is a first-order, time-varying, linear differential equation with a constant forcing term, the solution of which is given by

$$n(\tau) = \frac{\alpha\sqrt{d}}{(\sqrt{d}\cosh\sqrt{d}\,(T-\tau) - a\sinh\sqrt{d}\,(T-\tau))}\left[\frac{a}{\sqrt{d}}\sinh\sqrt{d}\,(T-\tau) - \cosh\sqrt{d}\,(T-\tau)+1\right] + \frac{\beta}{J\,(\sqrt{d}\cosh\sqrt{d}\,(T-\tau) - a\sinh\sqrt{d}\,(T-\tau)}\left[\sinh\sqrt{d}\,(T-\tau) - \frac{a}{\sqrt{d}}\cosh\sqrt{d}\,(T-\tau) + \frac{a}{\sqrt{d}}\right] \tag{25}$$

From Eqs. (14), (19), (24), and (25) the following feedback control solution is obtained:

$$u^{*}(t) = \left\{\frac{k_b}{2} + \frac{k_t a}{2Jb} + \left[\frac{1-\cosh\sqrt{d}\,(T-t)}{\sinh\sqrt{d}\,(T-t)}\right]\left[\frac{k_t\sqrt{d}}{2Jb}\right]\right\} x(t) + \left\{\frac{k_t}{2J^2 b} + \left[\frac{1-\cosh\sqrt{d}\,(T-t)}{\sinh\sqrt{d}\,(T-t)}\right]\left[\frac{k_t a}{2J^2 b\sqrt{d}}\right]\right\}\beta + \left\{\left[\frac{1}{\sinh\sqrt{d}\,(T-t)}\right]\left[\frac{k_t\sqrt{d}}{2Jb}\right]\right\}\left[\alpha - x(t)\right] \tag{26}$$

where

$\eta(\tau)$ is replaced by $\lambda(t)$ and $z(\tau)$ is replaced by $x(t)$ without loss of generality. In Eq. (26) at $t = T$, using L'Hospital's rule, one finds:

$$u^{*}(T) = \left[\frac{k_b}{2} + \frac{k_t a}{2Jb}\right]x(T) + \left|\frac{k_t}{2J^2 b}\right|\beta + \left[\frac{k_t}{2Jb}\right]\dot{x}(T) \tag{27}$$

If, however, $\lim_{t\to T}\left[\frac{\alpha - x(t)}{\sinh\sqrt{d}\,(T-t)}\right] \equiv 0$ for $t \geq T$ until a new disturbance or a new desired speed-setting is applied to the system, then:

$$u^{*}(T) = \left[\frac{k_b}{2} + \frac{k_t a}{2J_e b}\right]x(T) + \left[\frac{k_t}{2J_e^{\,2} b}\right]\beta \tag{28}$$

Substituting Eq. (28) into Eq. (8) and using definitions of a and b given before yields:

$$\dot{x}(T) \equiv 0 \text{ for } t \geq T \tag{29}$$

If this requirement is met, the controller enters the steady-state phase of its operation. Therefore, it is clear that some means must be incorporated into the control system to turn the time-varying feedback gains on at t = 0 and to turn them off at t = T. Note that in Eq. (26):

$$-1 \leq \left[\frac{1 - \cosh\sqrt{d}\quad(T-t)}{\sinh\sqrt{d}\quad(T-t)}\right] \leq 0 \text{ in } 0 \leq t \leq T \tag{30}$$

$$\lim_{t \to T}\left(\frac{1}{\sinh\sqrt{d}\quad(T-t)}\right) = \infty \tag{31}$$

If, and only if, $\sqrt{d}\,T$ is sufficiently small, it seems reasonable to replace the optimal control law as given in Eq. (26) by a suboptimal control law of the following form:

$$u^*_{so}(t) = \left[\frac{k_b}{2} + \frac{k_t a}{2Jb}\right] x(t) + \left(\frac{k_t}{2J^2 b}\right) \beta + \left[\frac{k_t\sqrt{d}}{2Jb}\right] \left[\frac{1}{\sinh\sqrt{d}\quad(T-t)}\right]\left[\alpha - x(t)\right] \tag{32}$$

As a result of this further simplification, the practical implementation of the control law is greatly simplified without significantly degrading the optimality of the system as will be shown later.

Substituting $u^*_{so}(t)$ for u(t) in Eq. (8) yields:

$$x(t) = x(0)\left[\frac{\cosh\sqrt{d}\quad T+1}{\cosh\sqrt{d}\quad T-1}\right]^{\frac{1}{2}}\left[\frac{\cosh\sqrt{d}\quad(T-t)-1}{\cosh\sqrt{d}\quad(T-t)+1}\right]^{\frac{1}{2}} + \alpha\left\{1 - \left[\cosh\sqrt{d}\quad(T-t)-1\right]^{\frac{1}{2}} \cdot \left[\frac{\sinh\sqrt{d}\quad T}{\left[(\cosh\sqrt{d}\quad(T-t)+1\right]^{\frac{1}{2}}\left[\cosh\sqrt{d}\quad(T-1)\right]}\right]\right\} \tag{33}$$

From Eq. (33) it is clear that:

$$x(t)\ \Big|_{t=0} = x(0) \tag{34}$$

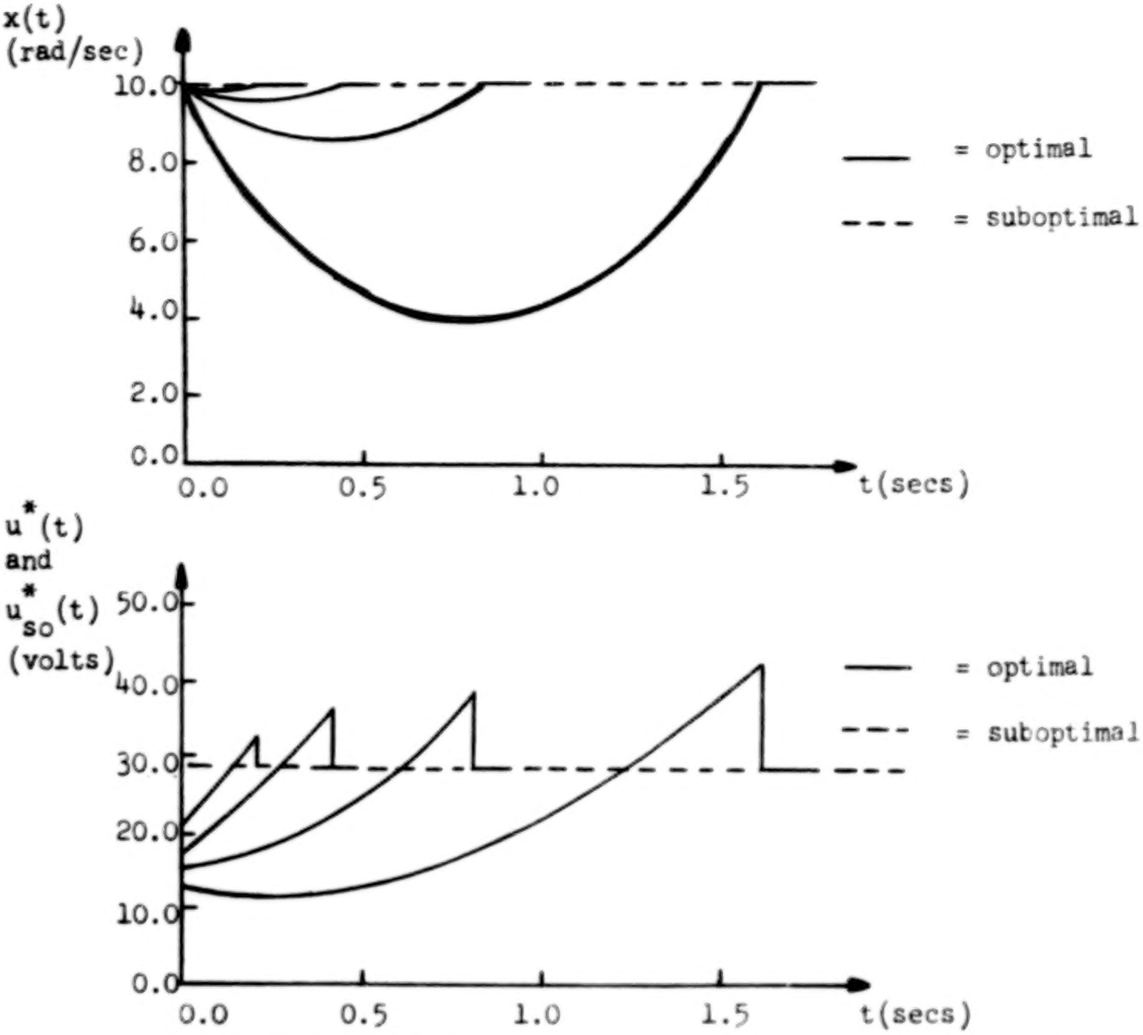
$$x(t)\ \Big|_{t=T} = \alpha \text{ as required in any control action} \tag{35}$$

The salient features of the optimal and suboptimal control systems are obtained from appropriate computer simulations by using Eqs. (8), (9), (26), and (32). Some of the results are shown in Figures 2, 3, and 4. The following remarks are made to supplement the results:

1. For the case of speed-control, i.e., Figures 2 and 3: $\dot{x}(t) = 0$ at $t = \frac{T}{2}$ for all α's, β's, and T's. The optimal control $u^*(t)$, hence the optimal trajectory $x(t)$, is greatly influenced by the choice of T. For $u^*(t)$, E approximately changes by the same factor as the response time T. The suboptimal control $u^*_{so}(t)$, hence $x(t)$, is constant for any T. For small T's the performance characteristics of optimal and suboptimal control systems are almost identical. Figure 2 shows that more energy is supplied from the battery to the motor for the case of suboptimal system than that of the optimal system. Figure 3 shows that more energy is supplied from the motor to the battery for the case of optimal system than that of the suboptimal system.

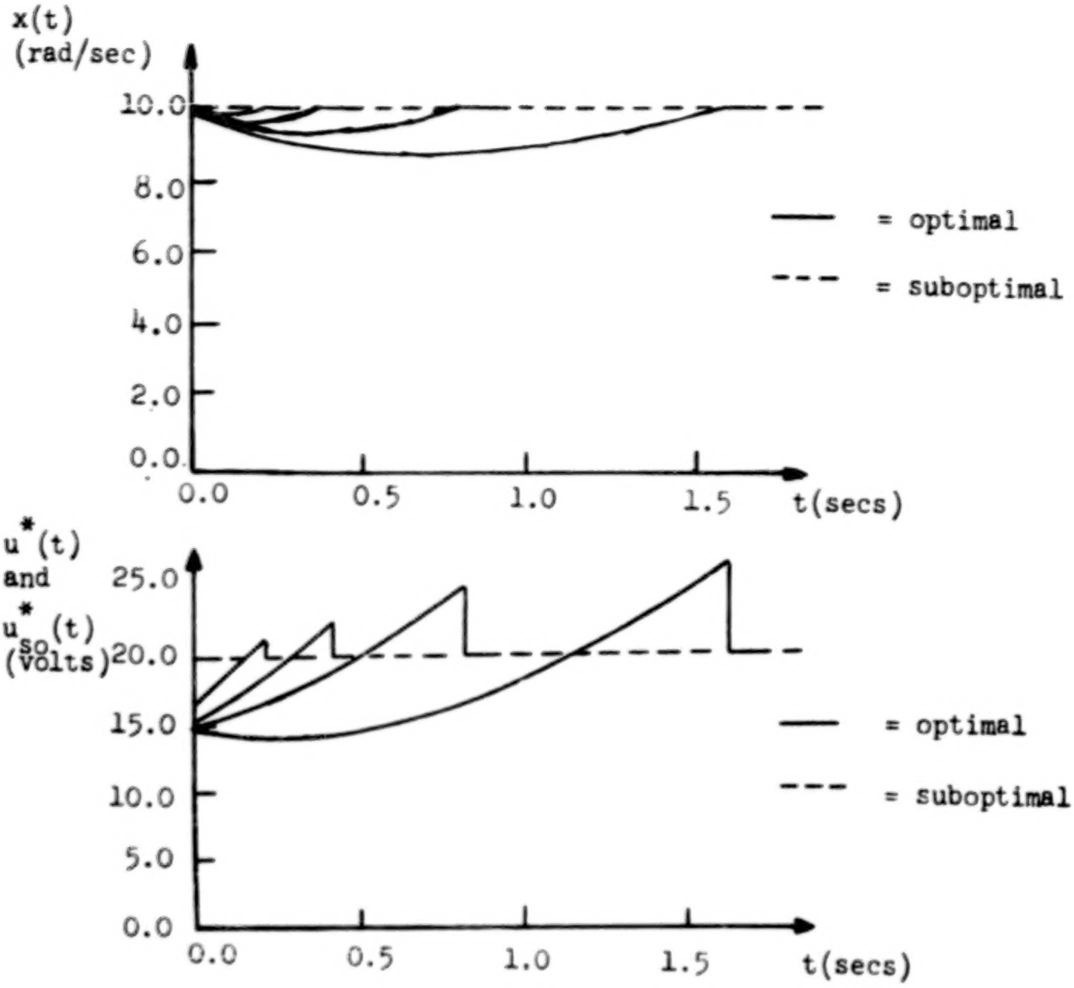
2. For the case of speed-setting, i.e., Figure 4:

   $u^*(t)$, $u^*_{so}(t)$, and $x(t)$ change linearly with time in $0 \leq t \leq T$ for small T's. Note, however, small T means a high acceleration requirement which, in turn, demands a high control voltage $u(t)$, and from Eq. (9) it is clear that the control energy consumption becomes very large. This is because the integrand in Eq. (9) is dominated by $u^2(t)$ and the integral of this quantity over $(0,T)$ is extremely large even though T is small. The control energy consumption E becomes independent of T for large T's. This is because energy saved by keeping the control voltage at a low value for most portions of $(0,T)$, i.e., low acceleration is diminished by requiring higher acceleration as $t \to T$ for large T's. Therefore, the net gain in control energy consumption for T greater than a certain value is extremely negligible.



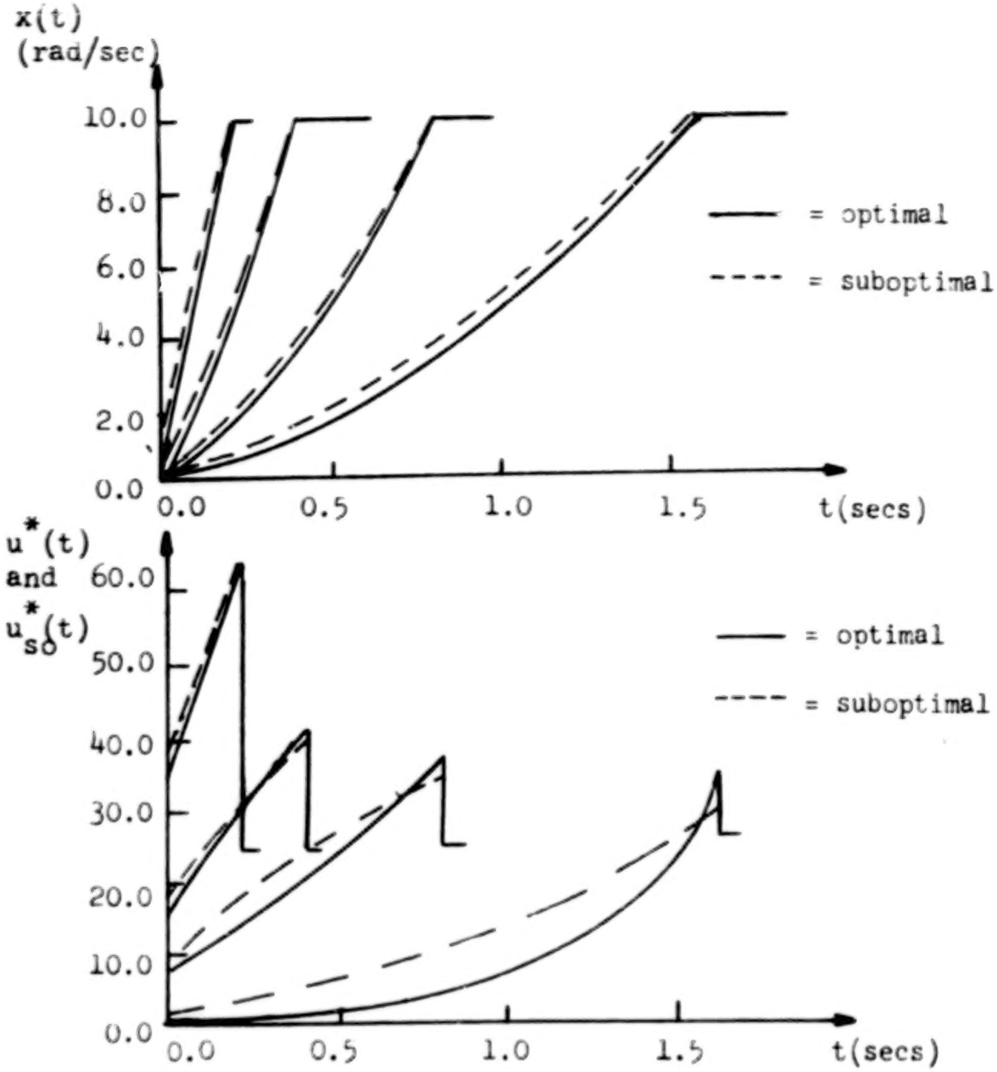
T (sec)	ENERGY CONSUMPTION	
	Optimal (watt-seconds)	Suboptimal (watt-seconds)
0.2	52.59	53.20
0.4	102.62	106.40
0.8	188.82	212.80
1.6	286.38	425.60

Figure 2.- Results for: α = 10.0 rad/sec, v(0-) = 0.0 newton-meter, β = 10.0 newton-meter, x(0) = 10.0 rad/sec (speed-control), u*(0) = 24.2 volts



T (sec)	Energy Consumption	
	Optimal (watt-seconds)	Suboptimal (watt-seconds)
0.2	- 3.372	- 3.0
0.4	- 6.8372	- 6.0
0.8	-14.23	- 12.0
1.6	-31.21	- 24.0

Figure 3.- Results for: α = 10.0 rad/sec
$v(0-)$ = 0.0 newton-meter,
β = -10.0 newton-meter,
$x(0)$ = 10.0 rad/sec (speed control),
$u^*(0)$ = 24.2 volts



T (sec)	Energy Consumption	
	Optimal (watt-secs)	Suboptimal (watt-secs)
0.2	359.0	359.0
0.4	237.80	238.0
0.8	185.66	186.84
1.6	168.84	173.43

Figure 4.- Results for: α = 10.0 rad/sec, v(0-) = 0.0 newton-meter, β = 0.0 newton-meter, x(0) = -10.0 rad/sec (acceleration), u*(0-) = 0.0 volt

If it is required to decelerate the speed of the vehicle to a new desired setting, it has been determined, but not shown here that, a considerable amount of energy can be saved for small T's. On the other hand for large T's the energy which can be gained by keeping T small is lost by requiring very high accelerations as $t \to T$ for large T's.

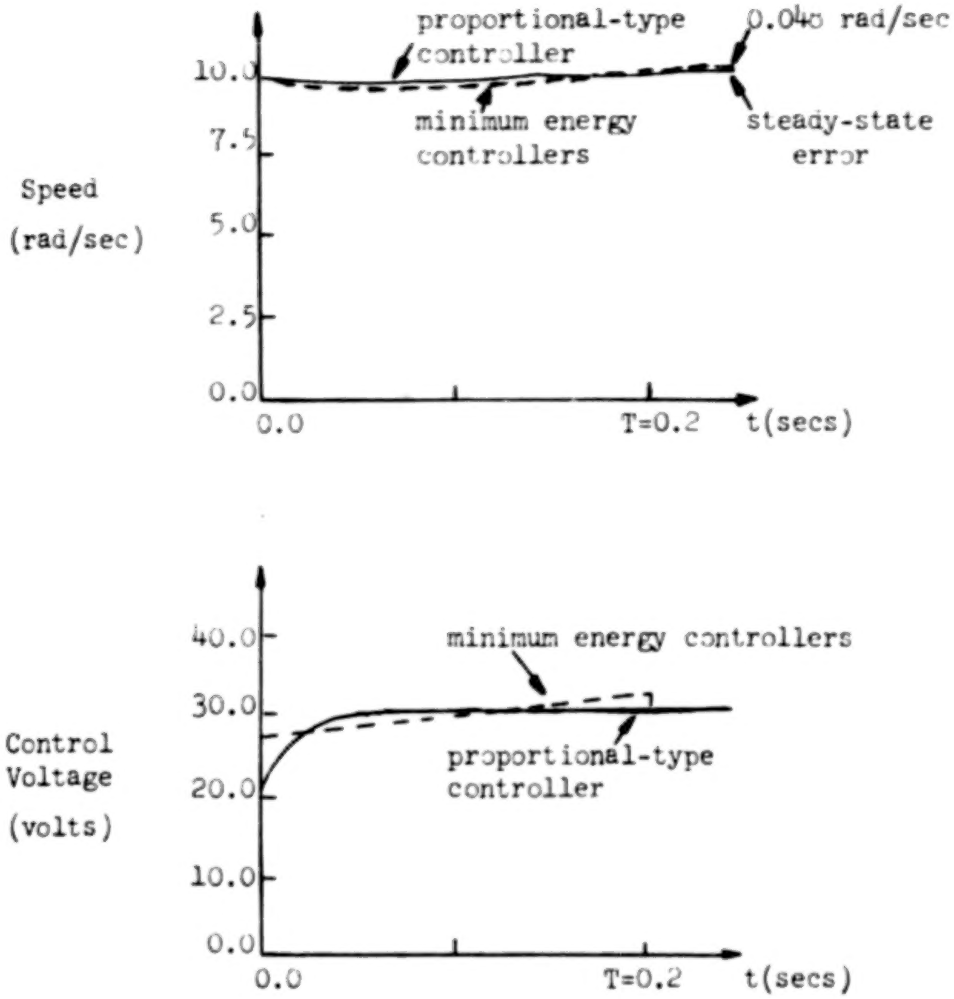
The performance of minimum-energy controllers, which are represented by Eqs. (26) and (32), are compared with the performances of the classical controllers of the following types under the assumption of zero armature inductance:

1. A controller with a high-gain amplifier in the forward path, which continuously operates on the error signal to produce the required control signal which, in turn, matches the output speed with the desired speed within a permissible error is called a proportional-type speed-controller.
2. A controller with an integrating amplifier in the forward path, which continuously integrates the error signal to produce the required control signal which, in turn , matches the output speed with the desired speed exactly is called an integral-type speed-controller.

The results are summarized in Figures 5 through 8. In view of these results it can be concluded that the performances of the minimum energy controllers show a marked superiority over the performance of classical controllers during the acceleration (and deceleration) of the vehicle. Figure 9 shows a comparison between the performance characteristics of the bang-bang controller and the minimum energy controllers applied to the plant described in Eqs. (1) and (2). Note that in reference (1) it is shown that the optimal control law consists of only the bang-bang control in (0,T) for the case of deceleration.

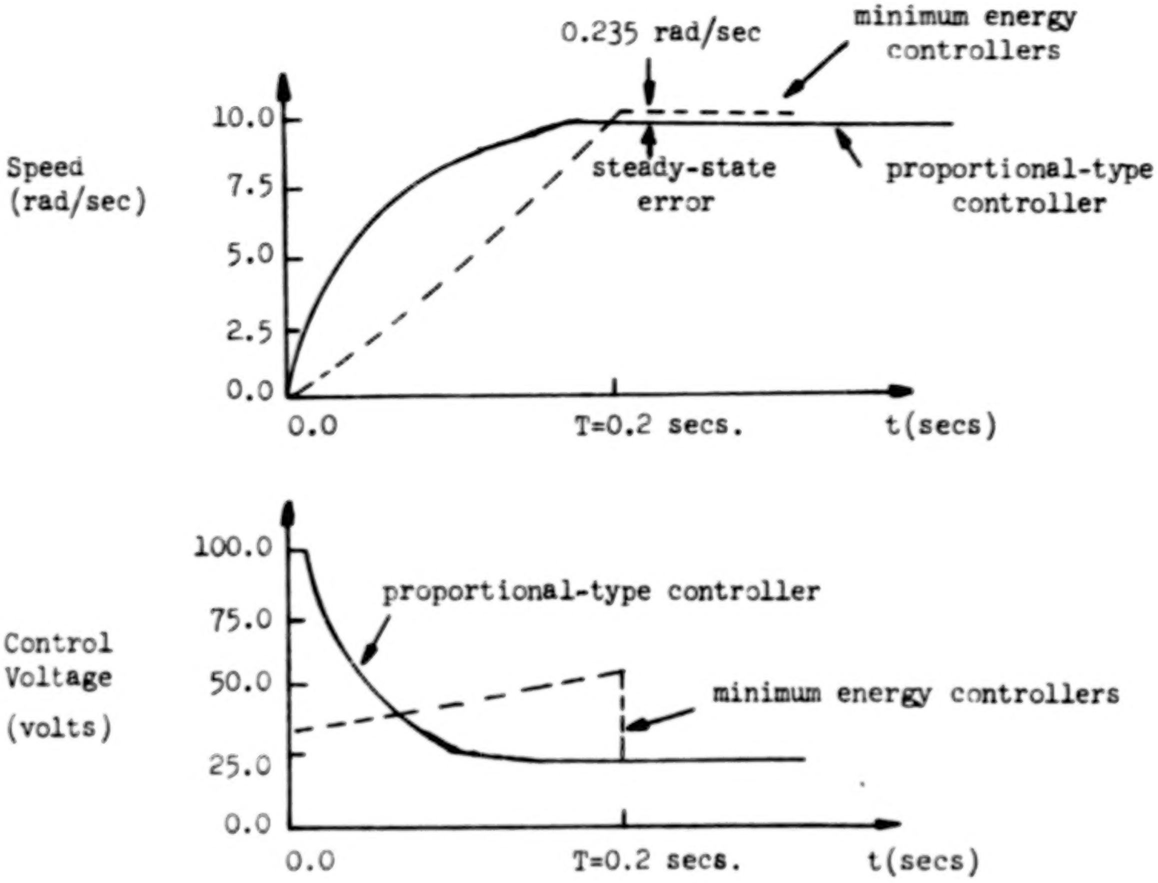
The two interesting features of the approximate optimization problem, which is obtained from the basic optimization problem, are as follows:

1. There is an infinite number of combinations between α and β and hence there is an infinite number of steady-state control voltages which must be specified by the bang-bang controller. The practical realization of such a controller is extremely complex and expensive.



T (sec)	Energy Consumption (watt-seconds)		
	Classical Controller	Optimal Controller	Suboptimal Controller
0.2	52.60	52.59	53.20

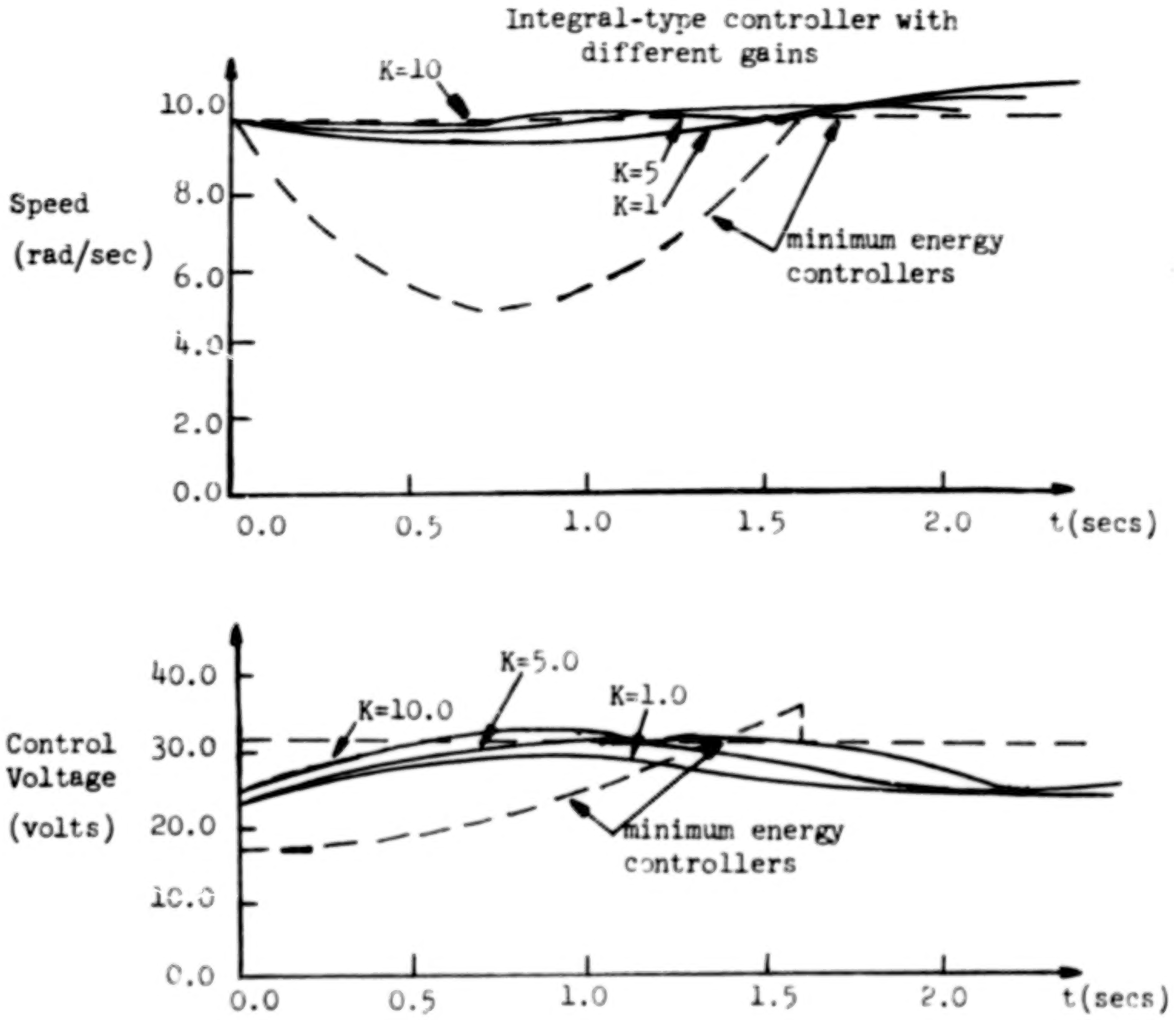
Figure 5.- Results for: $\alpha = 10.0$ rad/sec, $v(0-) = 0.0$ newton-meter, $\beta = 10.0$ newton-meter, $K = 100.0$



T (sec)	Energy Consumption (watt-seconds)		
	Classical Controller	Optimal Controller	Suboptimal Controller
0.2	3474.0*	359.00	359.0

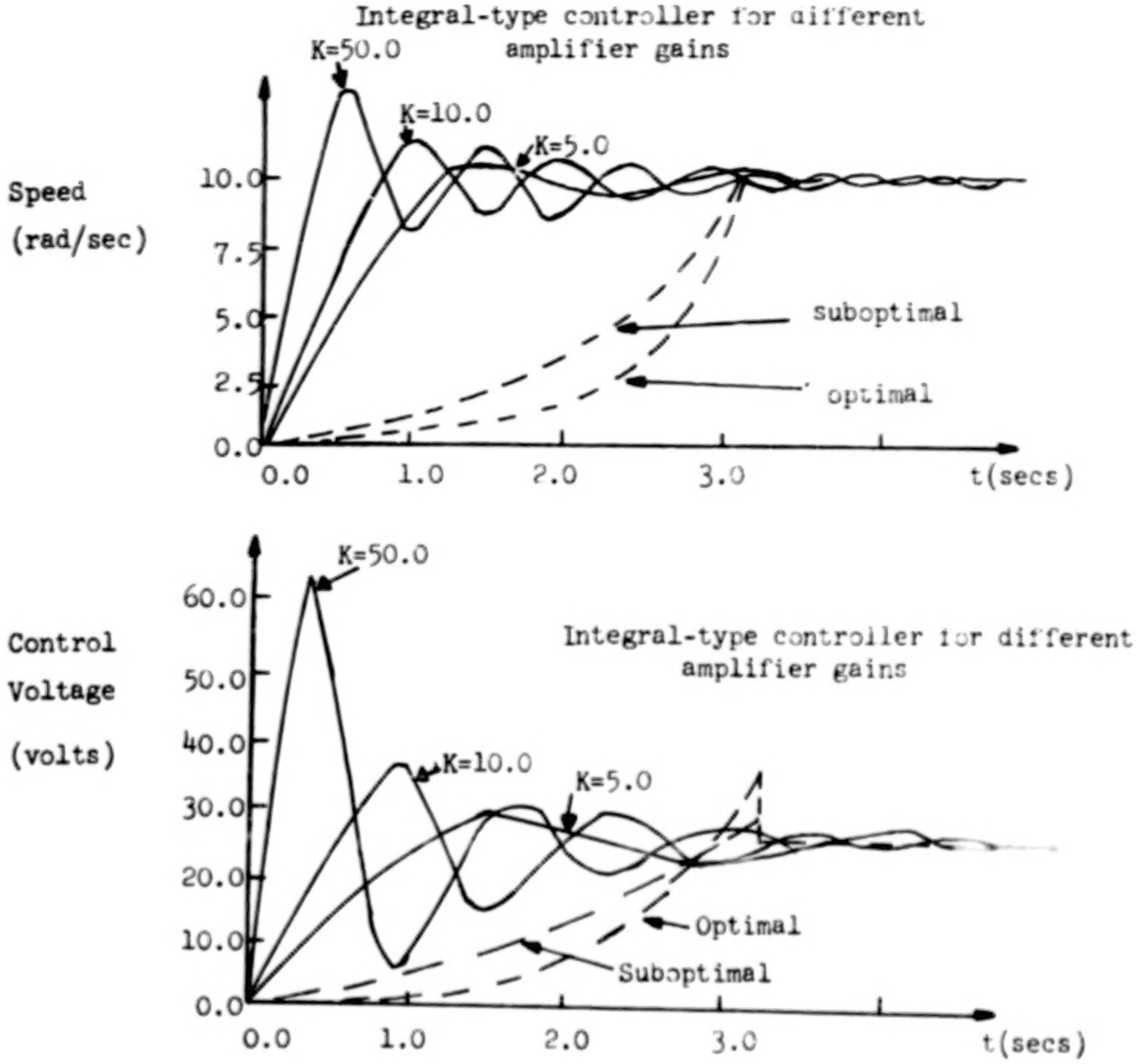
* Assuming control voltage is not saturated.

Figure 6.- Results for: α = 10.0 rad/sec,
$v(0-)$ = 0.0 newton-meter,
β = 0.0 newton-meter, K = 100.0



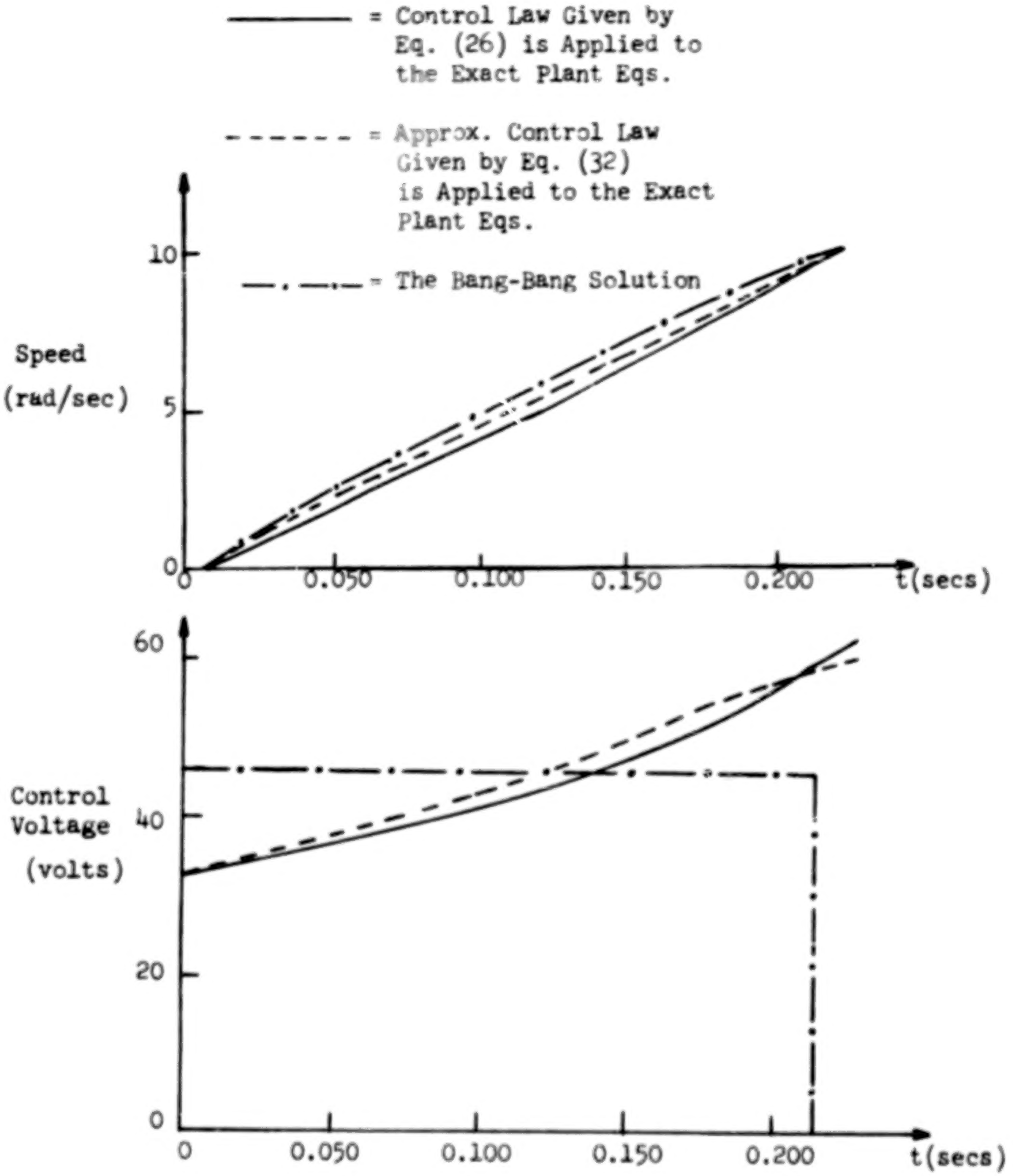
T (sec)	Energy Consumption (watt-seconds)		
	Classical Controller	Optimal Controller	Suboptimal Controller
1.6	414.36 (K=10) 392.74 (K=5) 368.00 (K=1)	286.38	425.00

Figure 7.- Results for: $\alpha = 10$ rad/sec,
$\beta = 10.0$ newton-meter



T (sec)	Energy Consumption (watt-secons)		
	Classical Controller	Optimal Controller	Suboptimal Controller
3.2	966.0 (K=50) 501.02 (K=10) 411.37 (K=5)	163.55	167.43

Figure 8.- Results for: α = 10 rad/sec, β = 0.0 newton-meter



	Energy Consumption (watt-seconds)		
T (sec)	Bang-Bang Solution	Exact Plant with Minimum Energy Controller	Exact Plant with Suboptimal Controller
0.225	334.0	344.42	343.85

Figure 9.- Comparsion of bang-bang solution with minimum-energy solutions for: α = 10.0 rad/sec, β = 0.0 newton-meter, and u = 45.0 volts

2. As shown in Figure 9, the performance characteristics of the minimum energy solutions of the approximate optimization problem are almost identical to those of bang-bang solution of the basic optimization problem.

Therefore, removing the inequality constraint on u(t) and neglecting motor armature inductance simplify the analysis considerably and results in the determination of an economical and practical control system.

THE STOCHASTIC OPTIMIZATION PROBLEM

In the above analysis it has been assumed that the disturbance torque remains constant during the controlling interval (0,T). In reality, the terrain profile encountered by the vehicles varies in a random fashion. Therefore, the disturbance torque is best described by a stochastic process. Other random disturbances acting on the system, such as terrain irregularities and wind gusts, are assumed to be negligible. The following stochastic optimization problem is originated from the approximate optimization problem for reasons of mathematical tractability and simplicity in the implementation of the control law:

Plant:

$$\dot{x}_1(t) = -\left[\frac{f}{J} + \frac{k_t k_b}{Jr}\right] x_1(t) - \left[\frac{f}{J} + \frac{k_t k_b}{Jr}\right]\alpha - \left[\frac{1}{J}\right] x_2(t)$$

$$- \left[\frac{1}{J}\right] x_3(t) + \left[\frac{k_t}{Jr}\right] u(t) \tag{36}$$

$$\dot{x}_2(t) = -(\omega_0)\, x_2(t) + \xi_w(t) \tag{37}$$

$$\dot{x}_3(t) = 0 \tag{38}$$

where

$$x_1(t) \triangleq x(t) - \alpha$$

$x_2(t) \triangleq$ a stationary Gaussian, exponentially correlated-noise process having zero mean, variance σ_c^2 and correlation time $1/\omega_0$. It can be shown that the $x_2(t)$ process, known as Ornstein-Uhlenbeck process (refs. 4-6), satisfies the Langevin equation given by Eq. (37).

Also, $\xi_w(t)$ is a stationary Gaussian, white-noise process having zero mean and spectral density $2\omega_0\sigma_c^2$.

$x_3(t) \triangleq$ constant disturbance torque, i.e., $x_3(t) \equiv \beta$ in the deterministic optimization problem studied above.

Performance Index:

$$E = F_w x_1^2(T) + \int_0^T \left\{ \frac{u^2(t) - k_b u(t)\ (x_1(t) + \alpha)}{r_a} \right\} dt \qquad (39)$$

where

T is fixed and u(t) is unconstrained.

$F_w \triangleq$ weighting factor which is artificially introduced here to facilitate the use of Bellman's dynamic programming technique (ref. 7).

The set of boundary conditions to be satisfied by the state vector

$\underline{x} = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$ for the three cases of control action are given below:

Case 1. Speed-Control.

$$\begin{aligned} x_1(0) &= 0 \quad , & x_1(T) &= 0.0 \\ x_2(0) &= 0 \quad , & x_2(T) &= \text{Free} \\ x_3(0-) \neq x_3(0) &= \beta \quad , & x_3(T) &= \beta \end{aligned} \qquad (40)$$

Case 2. Speed-setting.

$$\begin{aligned} x_1(0) &= x(0) - \alpha\,, & x_1(T) &= 0.0 \\ x_2(0) &= 0\,, & x_2(T) &= \text{Free} \\ x_3(0-) &= x_3(0) = \beta\,, & x_3(T) &= \beta \end{aligned} \tag{41}$$

Case 3. Speed-control and Speed-setting.

$$\begin{aligned} x_1(0) &= x(0) - \alpha\,, & x_1(T) &= 0.0 \\ x_2(0) &= 0\,, & x_2(T) &= \text{Free} \\ x_3(0-) &\neq x_3(0) = \beta\,, & x_3(T) &= \beta \end{aligned} \tag{42}$$

The solution of this stochastic optimization problem is obtained as follows:

Define the optimal expected value function for this optimization by

$$S(t,\underline{x}) = \underset{u(t)}{\text{Min}}\ \varepsilon_{\xi_w(t)} \left[F_w x_1^2(T) + \int_t^T \left\{ \frac{u^2(\tau) - k_b u(\tau)\ (x_1(\tau) + \alpha)}{r} \right\} d\tau \right] \tag{43}$$

where

$\varepsilon(\cdot) \triangleq$ expectation operator.

Applying the standard procedure (refs. 7,8) to Eq. (43) yields:

$$u^*(t) = \left[\left[\frac{k_b}{2} \right] (x_1(t) + \alpha) \right] - \left[\frac{k_t}{2J} \right] S_{x_1}(t,\underline{x}) \tag{44}$$

where

$$S_{x_1}(t,\underline{x}) = \frac{\partial}{\partial x_1} S(t,\underline{x})$$

and $S_{x_1}(t,\underline{x})$ is obtained by solving the following Bellman-Hamilton-Jacobi equation:

$$\begin{aligned} S_t(t,\underline{x}) &- \left(\tfrac{c}{2}\right)(x_1(t) + \alpha)^2 - a(x_1(t) + \alpha)S_{x_1}(t,\underline{x}) \\ &- \left(\tfrac{b}{2}\right)S^2_{x_1}(t,\underline{x}) - \left(\tfrac{1}{J}\right)x_2(t)S_{x_1}(t,\underline{x}) - \left(\tfrac{1}{J}\right)x_3(t)S_{x_1}(t,\underline{x}) \qquad (45) \\ &- (\omega_0)x_2(t)S_{x_2}(t,\underline{x}) + (\omega_0\sigma_c{}^2)S_{x_2x_2}(t,\underline{x}) = 0 \end{aligned}$$

where

$$S(T,\underline{x}) = F_w x_1{}^2(T) \qquad (46)$$

with the boundary condition:

a, b, and c are defined as before and

$$S_t(t,\underline{x}) = \frac{\partial}{\partial t}S(t,\underline{x})$$

$$S_{x_2}(t,\underline{x}) = \frac{\partial}{\partial x_2}S(t,\underline{x})$$

$$S_{x_2x_2}(t,\underline{x}) = \frac{\partial^2}{\partial x_2 \partial x_2}S(t,\underline{x})$$

The solution of Eq. (45) is obtained as follows:

Let:

$$S(t,\underline{x}) = \tfrac{1}{2} < \underline{x}, P(t) \ \underline{x} > + \ \underline{q} \ \underline{x} + h(t) \tag{47}$$

where

$P(t) \triangleq$ a 3x3 symmetric matrix

$<,> \triangleq$ inner product of two vectors

$\underline{q} \triangleq$ a raw vector with two components

$h(t) \triangleq$ a scalar function

Determining $s_t(t,\underline{x})$, $S_{x_1}(t,\underline{x})$, $S_{x_2}(t,\underline{x})$, $S_{x_2x_2}(t,\underline{x})$ from Eq. (47) and substituting them into Eq. (45), collecting the coefficients of equal powers of x's and equating them to zero yields the following equations:

$$\dot{P}_{11}(t) - bP_{11}^2(t) - 2aP_{11}(t) - c = 0 \tag{48}$$

$$P_{11}(T) = 2F_w \tag{49}$$

$$\dot{P}_{12}(t) - (a + bP_{11}(t) + \omega_0)P_{12}(t) - (\tfrac{1}{J})P_{11}(t) = 0 \tag{50}$$

$$P_{12}(T) = 0 \tag{51}$$

$$\dot{P}_{13}(t) - (a + bP_{11}(t))P_{13}(t) - (\tfrac{1}{J})P_{11}(t) = 0 \tag{52}$$

$$P_{13}(T) = 0 \tag{53}$$

$$\dot{P}_{22}(t) - (2\omega_0)P_{22}(t) - bP^2_{12}(t) - (\tfrac{2}{J})P_{12}(t) = 0 \tag{54}$$

$$P_{22}(T) = 0 \tag{55}$$

$$\dot{P}_{23}(t) - \omega_0 P_{23}(t) - bP_{12}(t)P_{13}(t) - (\frac{1}{J})(P_{12}(t) + P_{13}(t)) = 0 \tag{56}$$

$$P_{23}(T) = 0 \tag{57}$$

$$\dot{P}_{33}(t) - bP_{13}^2(t) - (\frac{2}{J})P_{13}(t) = 0 \tag{58}$$

$$P_{13}(T) = 0 \tag{59}$$

$$\dot{q}_1(t) - (a + bP_{11}(t))\, q_1(t) - (a\alpha)P_{11}(t) - \alpha c = 0 \tag{60}$$

$$q_1(T) = 0 \tag{61}$$

$$\dot{q}_2(t) - \omega_0 q_2(t) - (\frac{1}{J} + bP_{12}(t))\, q_1(t) - (a\alpha)P_{12}(t) = 0 \tag{62}$$

$$q_2(T) = 0 \tag{63}$$

$$\dot{q}_3(t) - a\alpha P_{13}(t) - bP_{13}(t)\, q_1(t) - (\frac{1}{J})q_1(t) = 0 \tag{64}$$

$$q_3(T) = 0 \tag{65}$$

$$\dot{h}(t) - (\frac{b}{2})q_1^2(t) - (a\alpha)q_1(t) - (\frac{c}{2})\alpha^2 + (\omega_0 \sigma_c{}^2)P_{22}(t) = 0 \tag{66}$$

$$h(T) = 0 \tag{67}$$

Since:

$$S_{x_1}(t,\underline{x}) = P_{11}(t)\, x_1(t) + P_{12}(t)\, x_2(t) + P_{13}(t)\, x_3(t) + q_1(t) \tag{68}$$

where

$P_{11}(t)$, $P_{12}(t)$, $P_{13}(t)$ and $q_1(t)$ can be determined in this case by solving analytically Eqs. (48), (50), (52) and (60) and using their corresponding boundary conditions results in:

$$P_{11}(t) = -\left\{\frac{a}{b} - \left(\frac{\sqrt{d}}{b}\right)\left[\frac{\cosh\sqrt{d}\,(T-t)}{\sinh\sqrt{d}\,(T-t)}\right]\right\} \tag{69}$$

$$P_{12}(t) = \left[\left(\frac{1}{Jb(\omega_0{}^2-d)}\right)\right]\left\{a\omega_0 + d + \sqrt{d}(a+\omega_0)\left[\frac{e^{-\omega_0(T-t)} - \cosh\sqrt{d}\,(T-t)}{\sinh\sqrt{d}\,(T-t)}\right]\right\} \tag{70}$$

Note that $\omega_0^2 = d$ does not represent a singularity in Eq. (70) since $\lim\limits_{\omega_0^2 \to d} P_{12}(t)$ exists.

$$P_{13}(t) = -\left(\frac{1}{Jb\sinh\sqrt{d}\,(T-t)}\right)\left\{\sinh\sqrt{d}\,(T-t) - \frac{a}{\sqrt{d}}(\cosh\sqrt{d}\,(T-t)-1)\right\} \tag{71}$$

$$q_1(t) = \left(\frac{-\alpha\sqrt{d}}{b\sinh\sqrt{d}\,(T-t)}\right)\left\{\frac{a}{\sqrt{d}}\sinh\sqrt{d}\,(T-t) - \cosh\sqrt{d}\,(T-t) + 1\right\} \tag{72}$$

where

$$d \triangleq a^2 - bc$$

and the operation $\lim F_w \to \infty$ has been performed on the expressions obtained from the solutions of Eqs. (48), (50), (52) and (60). Substituting Eq. (68) into Eq. (44) yields:

$$\begin{aligned} u^*(t) = {} & \left\{\frac{k_b}{2} + \frac{k_t a}{2Jb} + \left[\frac{1-\cosh\sqrt{d}\,(T-t)}{\sinh\sqrt{d}\,(T-t)}\right]\left[\frac{k_t\sqrt{d}}{2Jb}\right]\right\}x(t) + \left\{\frac{k_t}{2J^2b} + \right. \\ & \left.\left[\frac{1-\cosh\sqrt{d}\,(T-t)}{\sinh\sqrt{d}\,(T-t)}\right]\left[\frac{k_t a}{2J^2b\sqrt{d}}\right]\right\}x_3(t) + \left\{\left[\frac{1}{\sinh\sqrt{d}\,(T-t)}\,\frac{k_t\sqrt{d}}{2Jb}\right]\right\}(\alpha - x(t) \\ & -\left\{aw_o + d + \left[\frac{e^{-\omega_0(T-t)} - \cosh\sqrt{d}\,(T-t)}{\sinh\sqrt{d}\,(T-t)}\right]\left[\sqrt{d}(a+\omega_0\right]\right\}\left[\frac{k_t}{2J^2b(\omega_0{}^2-d)}\right]x_2(t) \end{aligned} \tag{73}$$

In Eq. (73) making the following approximations:

$$\left[\frac{1 - \cosh\sqrt{d}\ (T-t)}{\sinh\sqrt{d}\ (T-t)}\right] \equiv \lim_{t\to T}\ (\cdot) = 0$$

$$\{a\omega_0 + d + \sqrt{d}\ (a + \omega_0)\left(\frac{e^{-\omega_0(T-t)} - \cosh\sqrt{d}\ (T-t)}{\sinh\sqrt{d}\ (T-t)}\right)\} \equiv \lim_{t\to T} \{\cdot\}$$

$$= -\ (\omega_0^{\ 2} - d)$$

yields the following suboptimal solution for the stochastic control problem:

$$u^*_{so}(t) = \left[\frac{k_b}{2} + \frac{k_t a}{2Jb}\right]x(t) + \left[\frac{k_t}{2J^2 b}\right]\left[x_2(t) + x_3(t)\right] + \left[\frac{k_t\sqrt{d}}{2Jb}\right]\left(\frac{1}{\sinh\sqrt{d}\ (T-t)}\right)\left[\alpha - x(t)\right] \qquad (74)$$

The result given by Eq. (74) is important for the following reasons: In practice it is much easier to measure the disturbance torque v(t) which is the sum of state variables $x_2(t)$ and $x_3(t)$ rather than measuring them separately. Hence the implementation of the control law given by Eq. (74) is straightforward. The implementation of the control law given by Eq. (73) cannot be accomplished unless a Kalman filter (ref. 9) is used to estimate the state vector $\underline{x}$ from the measurable states. This follows from the concept of the separation principle (ref. 10) which allows the estimation of the state vector $\underline{x}$ and the computation of the control law to be performed independently (ref. 11).

Several interesting salient features of the stochastic optimal control system are obtained in reference 2. In summary, it is shown that

1. Speed variance which is defined by $\varepsilon\left\{\left[x(t) - x_d(t)\right]^2\right\}$ where

   $x(t) \triangleq$ stochastic speed

   $x_d(t) \triangleq$ deterministic speed

becomes zero at $t = T$. Hence $x(t) = \alpha$ as desired.

2. The control variance which is defined by
$\varepsilon \left\{[u^*(t) - u_d^*(t)]^2\right\}$
where

$$u^*(t) \overset{\Delta}{=} \text{stochastic control law}$$

$$u_d^*(t) \overset{\Delta}{=} \text{deterministic control law}$$

is bounded at $t = T$. This is because in Eqs. (73) and (74) the error term $[\alpha - x(t)]$, which is a stochastic process, decreases to zero fast enough as
$\lim_{t \to T} \left(\frac{1}{\sin\sqrt{d}\,(T-t)}\right) = \infty$ so that $u^*(t)$ is bounded at $t = T$. It can be shown that if $x_2(t)$ is assumed to be a stationary Gaussian white-noise with zero-mean, the control variance becomes unbounded at $t = T$.

3. When the stochastic optimal and suboptimal control laws are applied to the exact plant equations given by Eqs. (1) and (2) using the appropriate Monte Carlo simulation technique (ref. 2), the corresponding performance characteristics appear to be almost identical for certain T's.

It is now clear that Eq. (74) represents the control law which is to be implemented. Substituting the values of a,b,c, and d previously defined yields the following equation:

$$u^*(t) = \left[k_b + \frac{rf}{k_t}\right] x(t) + \left(\frac{r}{k_t}\right) v(t) + \left[\frac{fr}{k_t}\left[1 + \frac{k_t k_b}{fr}\right]^{\frac{1}{2}}\right]$$

$$\left[\frac{1}{\sinh\left(\frac{f}{J}\left[1 + \frac{k_t k_b}{fr}\right]^{\frac{1}{2}}[T - t]\right)}\right] \cdot [\alpha - x(t)] \tag{75}$$

where

$$u^*(t) \equiv u_{so}^*(t) \text{ without loss of generality.}$$

The control law given by Eq. (75) is a linear combination of feedback signals from a speed sensor, i.e., a tacho-generator, a terrain slope measuring device with a bias signal for tire resistance at zero speed, i.e., a pendulum-type device and a feed-forward signal from a throttle mechanism, i.e., a potentiometer.

The functional relationship between the terrain slope $\phi(t)$ and the disturbance torque v(t) for a wheeled-vehicle with rubber tires is given by

$$v(t) = (w_V \sin \phi(t) + f_1) \left[\frac{r_w c_f}{k_G} \right] \tag{76}$$

where

w_V = total weight of the vehicle [lb]

$\phi(t)$ = slope angle of the terrain with respect to local horizontal [deg]

f_1 = tire resistance at zero speed [lb)

r_w = effective radius of the wheel-tire assembly [ft]

k_G = gear ratio

c_F = conversion factor from ft-lb to newton-meter.

CONTROL LAW IMPLEMENTATION

The control law given by Eq. (75) is mechanized in the laboratory and the resulting minimum energy control system block diagram is shown in Figure 10 where

$k_1 \triangleq$ tacho-generator voltage gain,

$k_2 \triangleq$ disturbance torque-measuring device voltage gain,

$k_A \triangleq$ dc to dc SCR converter voltage gain.

In the laboratory, the effects of terrain slope changes on the system performance is simulated by means of a second permanent-magnet dc motor which is directly coupled to the drive motor shaft. The armature voltage of this second motor is controlled by a rheostat.

The prototype model is shown in Figure 11. The preliminary test results on the performance characteristics of the minimum-

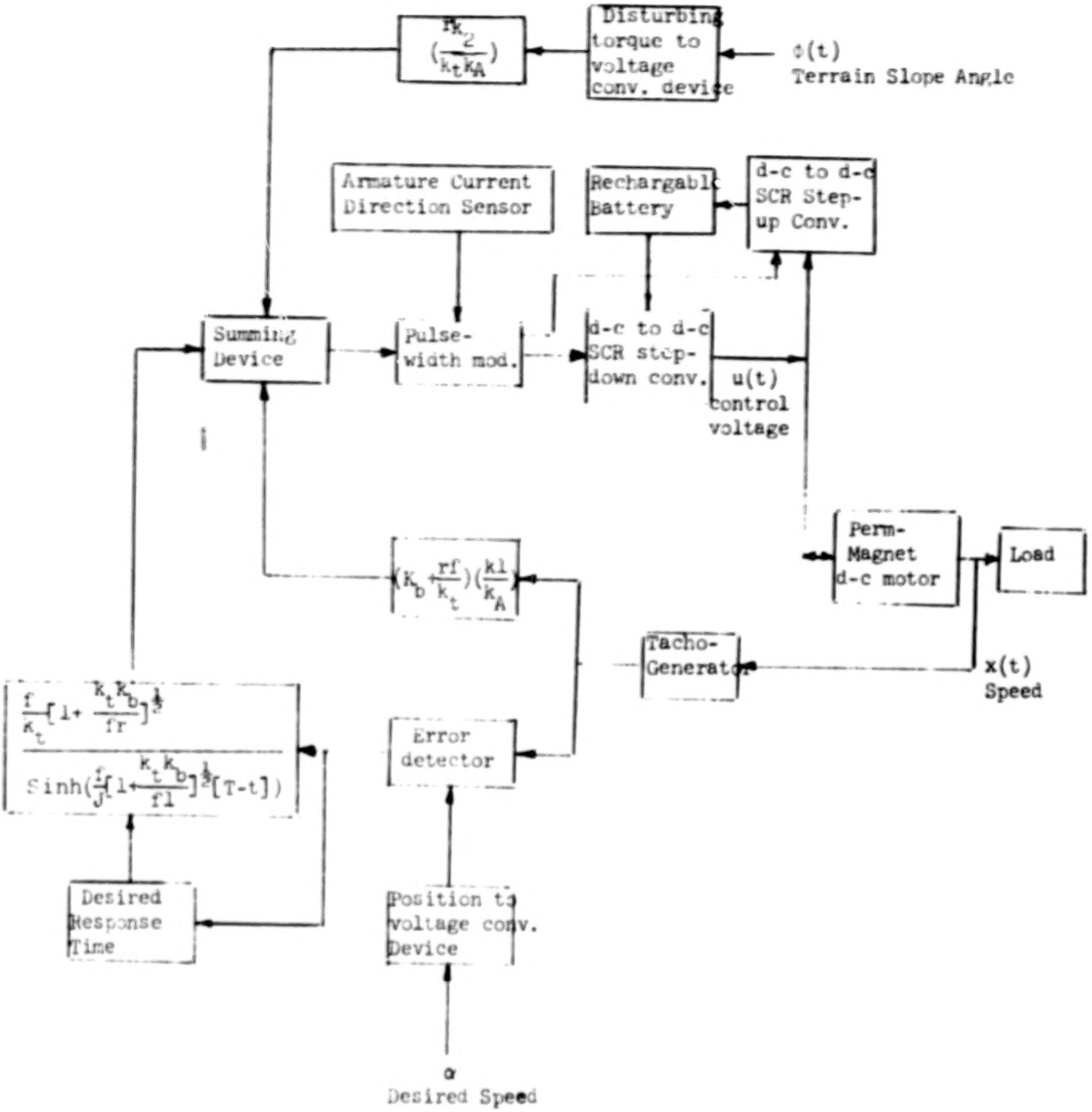

$(\frac{rk_2}{k_t k_A})$
Disturbing torque to voltage conv. device
$\phi(t)$
Terrain Slope Angle
Armature Current Direction Sensor
Rechargable Battery
d-c to d-c SCR Step-up Conv.
Summing Device
Pulse-width mod.
d-c to d-c SCR step-down conv.
u(t) control voltage
Perm-Magnet d-c motor
Load
$(K_b + \frac{rf}{k_t})(\frac{kl}{k_A})$
Tacho-Generator
x(t) Speed
$\frac{\frac{f}{K_t}[1+\frac{k_t k_b}{fr}]^{\frac{1}{2}}}{\sinh(\frac{f}{J}[1+\frac{k_t k_b}{fl}]^{\frac{1}{2}}[T-t])}$
Error detector
Desired Response Time
Position to voltage conv. Device
α
Desired Speed

Figure 10.- Minimum energy control system block diagram

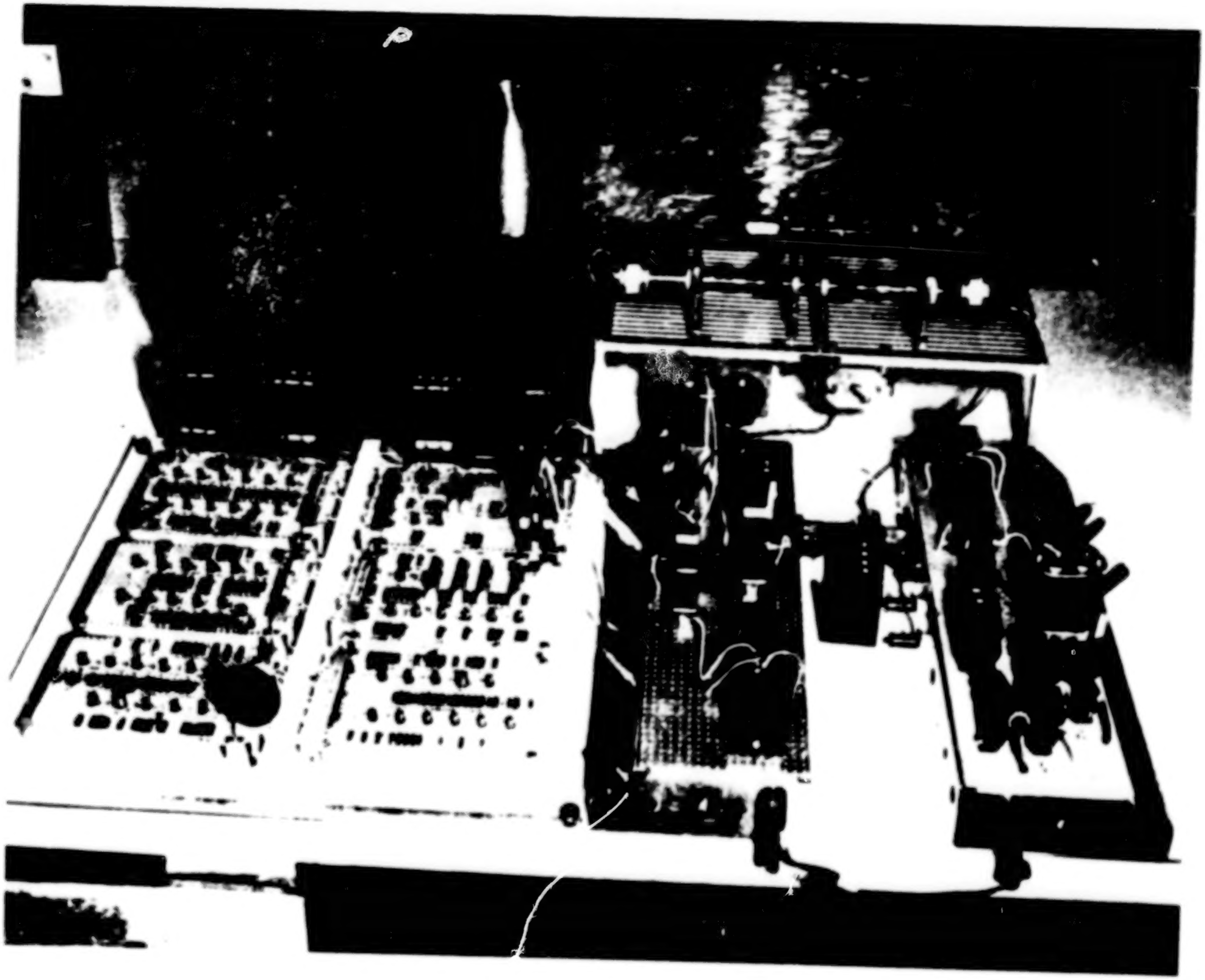


Figure 11.- Prototype model

energy controller, during the acceleration and speed-control actions show a very close relationship to theoretical results.

Some further testing is needed to verify the same relationships during the regenerative braking action. Note that the results presented here are equally applicable to the minimum energy control of dc motors utilizing separate field excitations.

CONCLUSIONS

The primary aim of this investigation has been to demonstrate the application of theoretical concepts of modern control theory into a class of engineering problems. Any application of an electric motor as a prime mover with electricity as the basic energy source in variable speed drive systems creates a need for a controller linking the energy source to the motor for speed setting and speed-control actions. Such systems span the applications spectrum from large industrial complexes to small portable devices. Recently an intensive effort initiated by several engineering institutions was made to extend them to electrically driven vehicles, such as electric cars, electric trains, electric earth-moving vehicles, electric lift trucks, submerged vehicles, and the like. However, as is typical of such problems solved by classical techniques, the existing designs are motivated largely by experience and in no sense are the resulting systems optimal. The recent advances in optimal control theory and digital computers make it feasible to replace all the existing empirical design techniques by a scientifically motivated and mathematically sound design procedure.

The minimum energy controller, which is mechanized in the laboratory, is applicable to fractional as well as integral horse power drive systems, provided that the necessary design changes are made in the dc to dc SCR step-up and step-down converters.

REFERENCES

1. Pontryagin, L.S.: The Mathematical Theory of Optimal Processes. Interscience Publishers, New York, New York, 1962.

2. Sahinkaya, Y.E.: Minimum Energy Control of Electric Propulsion Vehicles. Ph.D. Thesis, Electrical Engineering, California Institute of Technology, Pasadena, December 12, 1968.

3. Sridhar, R., Bellman, R.E., Kalaba, R.E.: "Sensitivity Analysis and Invariant Imbedding." RAND Corporation Memorandum, RM-4039-PR, 1964.

4. Uhlenbeck, G.E., and Ornstein, L.S.: "On the Theory of Brownian Motion." Selected Papers on Noise and Stochastic Processes, Nelson Wax (Editor), Dover, 1954, pp. 93-113.

5. Ash, G.: Optimal Guidance of Low-Thrust Interplanetary Space Vehicles. Ph.D. Thesis, Department of Electrical Engineering, California Institute of Technology, Pasadena, 1968.

6. Jordan, J.F.: Optimal Stochastic Control Theory Applied to Interplanetary Guidance. Ph.D. Thesis, University of Texas, Austin, 1966.

7. Dreyfus, S.E.: Dynamic Programming and the Calculus of Variations. Academic Press, 1965.

8. Sridhar, R.: Unpublished Class Notes. California Institute of Technology, Pasadena, 1966-67.

9. Kalman, R.E., and Bucy, R.: "New Results in Linear Filtering and Prediction Theory." Journal of Basic Engineering, 83D, March 1961, pp. 95-108.

10. Wonham, W.M.: "Stochastic Problems in Optimal Control." RIAS Report 63-14 (N65-16829), May 1963.

11. Bashein, G., Neuman, C.P., and Lavi, A.: " A Stochastic Perturbation Feedback Control with Applications." 1968 JACC Preprints, Univ. of Michigan, Ann Arbor, 1968.

Section 5

ADVANCED CONTROL TECHNIQUES

A NEW FORMULATION FOR THE EPSILON METHOD APPLIED TO THE MINIMUM-TIME-TO-CLIMB PROBLEM

By Lawrence W. Taylor, Jr., and Harriet J. Smith
Flight Research Center

SUMMARY

Balakrishnan's epsilon technique is used to compute minimum-time profiles for the F-104 airplane. This technique differs from the classical gradient method in that a quadratic penalty on the error in satisfying the equation of motion is included in the cost function to be minimized as a means of eliminating the requirement of satisfying the equations of motion. Although the number of unknown independent functions is increased to include the state variables, the evaluation of the gradient of the cost function is simplified, resulting in considerable computational savings, thereby making it appear feasible to use the epsilon method for real-time application.

There is considerable freedom in selecting the variables to be used in the epsilon method. The formulation discussed in this paper uses range and altitude as the variables. These variables are approximated by a functional expansion with unspecified coefficients which are determined by means of Newton's method. Typically, 8 to 10 iterations are required for convergence. A comparison is made of a solution obtained using the epsilon technique with a solution using the gradient method.

INTRODUCTION

Developing practical methods of solution to optimal control problems has been a continuing challenge to the mathematician as well as the engineer. One particular problem that has received considerable attention is that of determining the control function and corresponding airplane flight profile which minimizes the time in flying between specified flight conditions. Although a variety of techniques has been developed to obtain solutions to such problems, there remains a need for faster solutions.

Some of the methods that have been used in the past are the energy method (refs. 1, 2), dynamic programming (ref. 3), and the gradient method (refs. 4-9). Both the energy method and dynamic programming have the advantage of being fast; however, these methods are only approximate in the sense that the state is usually reduced to a single variable. The gradient method, on the other hand, gives accurate results but requires considerable computation time and storage. At NASA's Flight Research Center techniques that are fast enough to be used in real-time applications are being studied and developed. It was with this goal in mind that Balakrishnan's epsilon technique (ref. 10) was applied to the problem of computing optimum profiles for airplanes.

Balakrishnan (ref. 10) suggested adding a term to the cost function to be minimized, which penalizes departures from the equations of motion, and then treating the state as well as the control as unknown variables. The increase of unknown variables was claimed to be more than compensated for by the marked reduction in the amount of computation necessary to evaluate the gradient involved by virtue of their independence.

In reference 11 a detailed formulation of an application of the epsilon technique to the minimum-time-to-climb problem is presented. The numerical results for an F-4 airplane application using the epsilon method were compared with solutions obtained by a gradient technique and an energy method. Computation times were also presented and compared with those required by other methods of solution.

In reference 12 the epsilon method was applied to the F-104 airplane, and the approximate solution generated by the epsilon method was compared to an independently calculated profile for which the dynamic equations of motion were satisfied exactly but which used the control input of the epsilon solution.

The formulation of the epsilon technique which has been used thus far has been restricted to problems of minimizing the time with no consideration given to either range or changing weight. In an effort to increase the versatility of the epsilon method, it has been reformulated to include range and the effects of changing weight. In addition, the new formulation has reduced the number of unknown functions from three to two in the interest of further reducing the computational time required for solution.

This paper outlines the changes in the formulation, discusses the reasons for the changes, and presents comparisons with the gradient method and with flight results. In addition, the feasibility of using the epsilon method in a real-time flight-profile-optimization system is discussed. The F-104 airplane is used for all of the examples.

SYMBOLS

a_m	coefficient of functional expansion for range
b_m	coefficient of functional expansion for altitude
D	drag, lb
F	thrust, lb
f_m	functions used in functional expansion
g	gravitational constant, ft/sec^2
h	altitude, ft
J	cost, sec
L	lift, lb
m	function, index
$\underline{m}$	mass, slugs
N	number of time points
R	range, ft
T	total time, sec
t	time, sec
V	velocity, ft/sec
γ	flight-path angle, deg (radians)
Δ	increment
ε	weighting term in modified cost function
λ	relaxation factor
$\Delta(\cdot)$	gradient of $(\cdot)$

Subscripts:

min	minimum
N	final condition

o   nominal

Superscript:

T   transpose

A dot over a symbol indicates differentiation with respect to time.

DISCUSSION

The problem to be considered may be briefly stated as follows: Given the dynamic equations of motion of an airplane,

$$\ddot{R} = \frac{F - D}{\underline{m}} \cos \gamma - \frac{L}{\underline{m}} \sin \gamma$$

$$\ddot{h} = \frac{F - D}{\underline{m}} \sin \gamma + \frac{L}{\underline{m}} \cos \gamma - g$$

where

$$L = \underline{m}V\dot{\gamma} + \underline{m}g \sin \gamma$$

$$V = \sqrt{\dot{R}^2 + \dot{h}^2}$$

$$\gamma = \text{arc tan} \left(\frac{\dot{h}}{\dot{R}} \right)$$

and where F - D is a function of h, V and L, find the control input L(t) [or $\gamma(t)$] and the state R(t), h(t) which minimize the cost function J, which, for the example considered, equals the time (J = T) required to pass from the initial conditions, V_1, h_1, γ_1, to the final conditions V_N, h_N, and γ_N (V_N, h_N, or γ_N may be left free depending on the result desired), subject to the constraint of satisfying the dynamic equations of motion.

The gradient technique used the formulation of the problem as stated and evolved a solution using a recursive formula such as:

$$L(t)_{N+1} = L(t)_N + \gamma \Delta_L J$$

In the epsilon formulation, the cost is expanded to include a penalty for not satisfying the dynamic equations

$$J = T + \frac{1}{\varepsilon}\int_0^T \left(\ddot{R} - \frac{F - D}{\underline{m}} \cos \gamma + \frac{L}{\underline{m}} \sin \gamma\right)^2 dt + \frac{1}{\varepsilon}\int_0^T \left[\ddot{h} - \frac{F - D}{\underline{m}} \sin \gamma - \frac{L}{\underline{m}} \cos \gamma + g\right]^2 dt$$

and the constraint of satisfying the dynamic equations is removed. The unknown functions include the state as well as the control variables, which for the problem being considered might be altitude, velocity, and lift. However, there is considerable freedom in selecting what are to be the state and control variables. In the older epsilon formulation, altitude and veloctiy were used as state variables, but flight-patch angle was used as the control variable. In the newer formulation, the only variables used are range and altitude, and it is not clear which, if either, should be considered the control variable. At the other extreme, it would be possible to have altitude, range, velocity, lift, and flight-path angle as state and control variables for the same problem. It would appear that the most efficient formulation is the one with the fewest number of variables as long as the epsilon formulation of the cost function was used. This point needs to be studied to be certain, however.

The authors have chosen to represent the variables as functional expansions in the interest of providing an efficient and smooth representation. In the older formulation, linear functions were used to pass through the end conditions to which a summation of sine functions was added. The unknowns, then, were the coefficients of the functional expansions. In the newer formulation in terms of range and altitude, cubic functions are used to satisfy the end conditions to which are added functions which are products of sine functions. This is done in order to preserve the end conditions, which requires in this case functions, the derivatives of which are zero at the end points.

The equations for range and altitude are:

$$R = R_1 + \Delta R \left(\frac{3t^2}{T^2} - \frac{2t^3}{T^3} \right) + V_1 \cos \gamma_1 \left(t - \frac{2t^2}{T} + \frac{t^3}{T^2} \right)$$

$$+ V_N \cos \gamma_N - \frac{t^2}{T} + \frac{t^3}{T^2} + \sum a_m f_m$$

$$h = h_1 + \Delta h \left(\frac{3t^2}{T^2} - \frac{2t^3}{T^3} \right) + V_1 \sin \gamma_1 \left(t - \frac{2t^2}{T} + \frac{t^3}{T^2} \right)$$

$$+ V_N \sin \gamma_N \left(- \frac{t^2}{T} + \frac{t^3}{T^2} \right) + \sum b_m f_m$$

where

$$f_m = \sin \pi\frac{t}{T} \sin m\pi\frac{t}{T}$$

Because of the lack of success in using the gradient method of minimizing the epsilon form of the cost function, a modification of the Newton-Raphson method was used, and with considerable success. The modification consists of approximating the second gradient of the cost with respect to the unknown coefficients with the outer product of a first gradient. (Details are presented in references 11 and 12.) The result is that the considerable computation time and complexity of determining the second gradient required by the Newton-Raphson method are avoided.

RESULTS

The evaluation of a minimum-time profile is shown in Figure 1. The profile labeled initial is that which corresponds to the cubic functions used to satisfy the end conditions. It can be seen that after only three iterations the profile closely approximates that of the final trajectory. This is one of the advantages of the epsilon method; that is, it is not at all critical as to what initial profile is used. In using the gradient formulation, considerable effort is often required to obtain an initial profile that satisfies the end conditions and is sufficiently close to the optimum to insure convergence.

Figure 2 shows how rapidly the cost or equation error and time converge to their final values. The exact number of iterations will vary somewhat, depending upon the end conditions and the nature of the optimum profile. An example involving a zoom climb might require 8 to 10 iterations.

Because a functional expansion is used for each of the variables, one must investigate how many functions are required for an adequate representation. Figure 3 shows how the final values of the equation error and time depend on the number of functions used. It is also useful to look at independent calculations which satisfy the dynamic equations exactly, but which use the resulting flight-path angle in deciding the number of functions required. Eight to ten functions are usually sufficient.

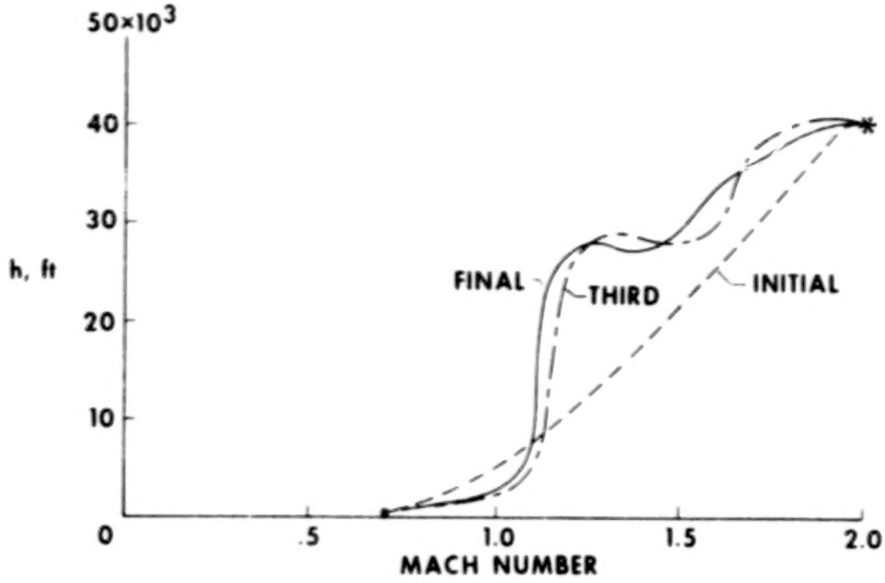


Figure 1.- Evolution of minimum time profile using epsilon technique

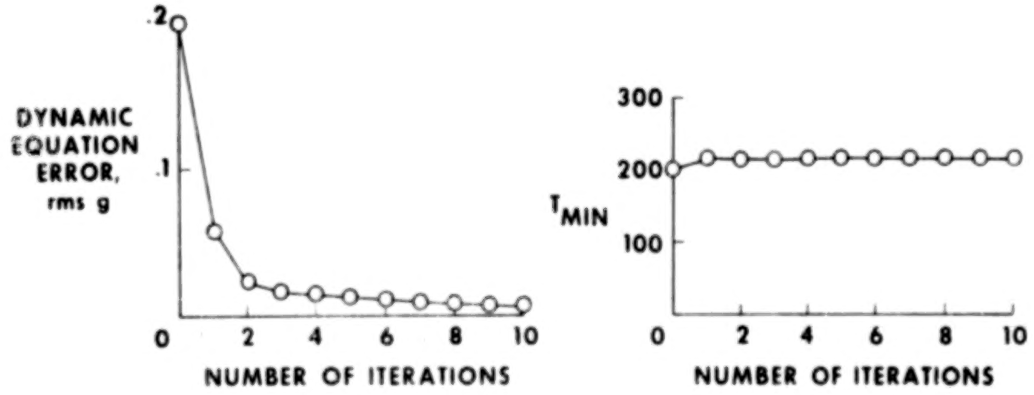


Figure 2.- Convergence of equation error and T_{MIN}

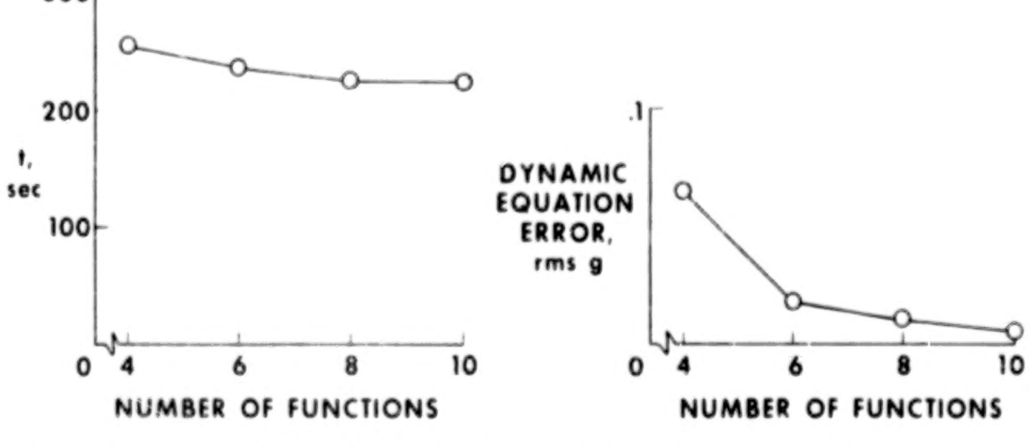


Figure 3.- Effect of number of functions on equation error and T_{MIN}

For the example being considered, the independently calculated profile in which the dynamic equations are satisfied exactly differed from the profile generated by the epsilon method at the terminal condition by 1 ft/sec and 1000 feet in velocity and altitude. When the comparison is that good, there is confidence in the validity of the results.

Figure 4 makes a comparison between the results obtained using the gradient and the epsilon methods. Although the differences between the profiles would appear to be significant, independent calculations show the corresponding times to be within a fraction of a percent. This is even more surprising because different sources of data were used. This result is further evidence that the epsilon method as formulated in terms of range and altitude gives valid results. A further check was made on the aerodynamic data tables by making a comparison between the time corresponding to the computed profile and the time required for an actual flight. The error was 11 percent, which is not considered large in view of the departures from the standard temperatures of 5° C.

It was stated at the beginning that interest at the Flight Research Center lies in developing a method that would enable optimum profiles to be computed in "real time." Figure 5 shows how this might work if the solution times were on the order of those obtained using the SDS 9300 digital computer. It is assumed that at a Mach number of 0.7 and an altitude of 500 feet the computer is instructed to compute a minimum-time profile to a Mach number of 2.0 and an altitude of 40,000 feet. Any terminal flight condition could have been chosen at this point. The airplane follows the profile given by the cubic functions which satisfy the end conditions until the first iteration has been generated. Nine seconds later the profile has been updated for the first time. This updated profile is then followed until there has been time for another update. Simultaneously, the fuel flow, air temperature, acceleration, and any other pertinent data could be used to enhance the accuracy of the aerodynamic, thrust, and atmospheric tables. In this way an optimum flight profile will be generated for an airplane having that particular performance and encountering that particular atmosphere. Is this really feasible? The authors believe it is. Although the SDS 9300 is considerably faster than available airborne computers, the solution times might prove to be comparable through the use of fixed-point arithmetic and through efficient programming.

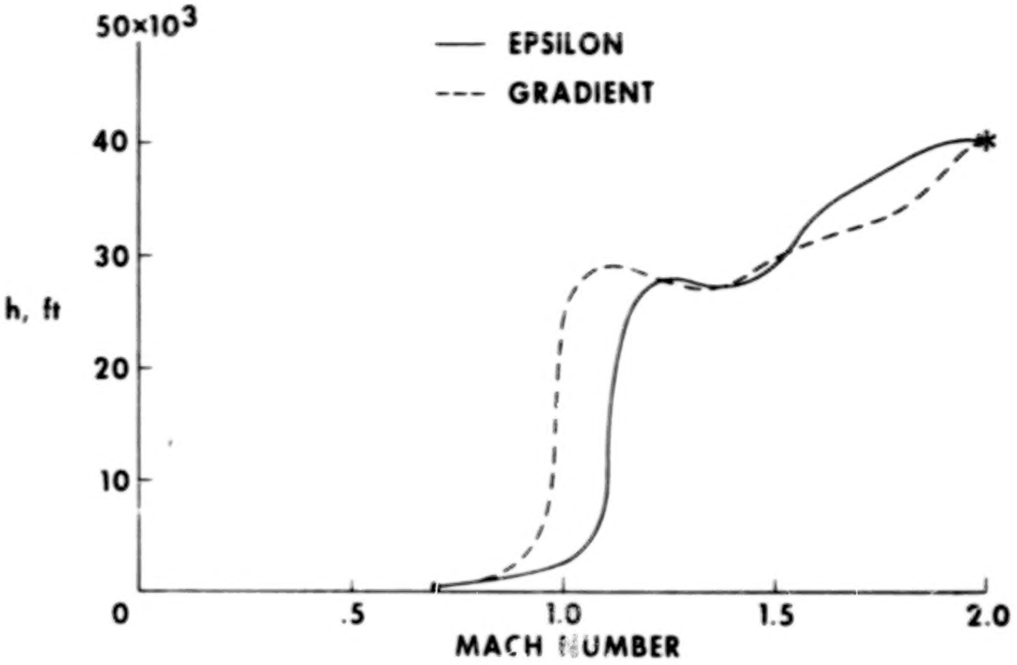


Figure 4.- Comparison of gradient and epsilon results

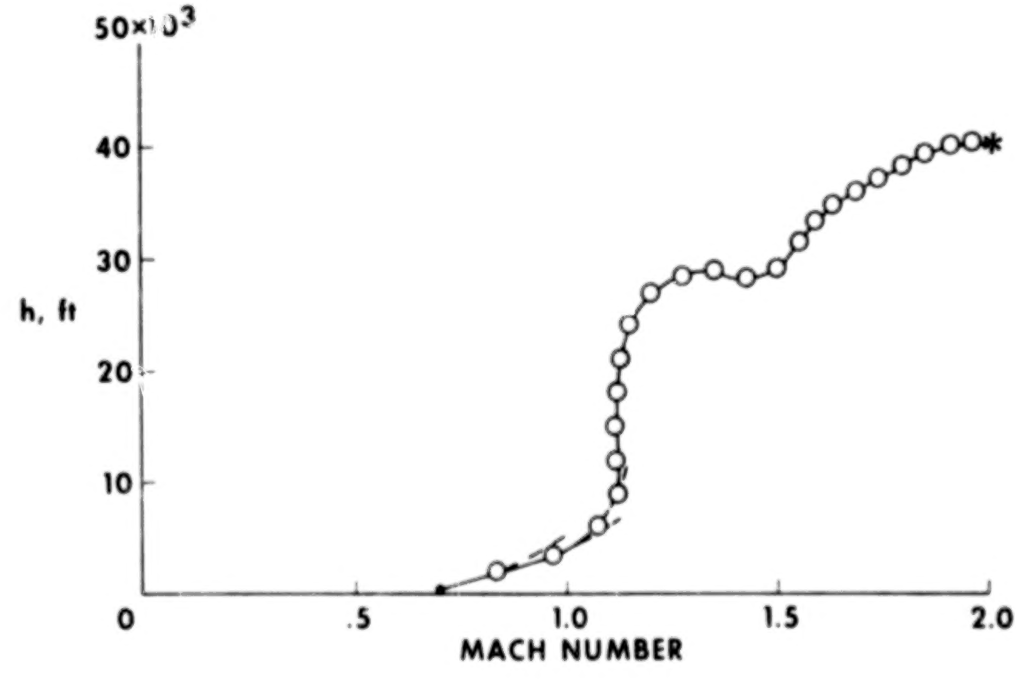


Figure 5.- A real-time application of the epsilon method

CONCLUDING REMARKS

The epsilon method as applied to the minimum-time-to-climb problem has been shown to give efficient and valid results, when compared to gradient results and actual flight.

It appears to be feasible to incorporate the epsilon method into a real-time optimum-profile system which would enable optimum flight to arbitrary terminal conditions while compensating for nonstandard performance and atmospheric characteristics.

REFERENCES

1. Lush, K. J.: "A Review of the Problem of Choosing a Climb Technique, With Proposals for a New Climb Technique for High Performance Aircraft." R&M 2257, British Aeronautical Research Council, 1951.

2. Rutowski, Edward S.: "Energy Approach to the General Aircraft Performance Problem." J. Aeron. Sci., vol. 21, March 1954, pp. 187-195.

3. Bellman, Richard; and Kalaba, Robert: Dynamic Programming and Modern Control Theory. Academic Press Inc., New York, c. 1965.

4. Kelley, H. J.: "Gradient Theory of Optimal Flight Paths." ARS J., vol. 30, no. 10, Oct. 1960, pp. 947-954.

5. Bryson, A. E.; and Denham, W. F.: "A Steepest-Ascent Method for Solving Optimum Programming Problems." J. Applied Mech., vol. 29, June 1962, pp. 247-257.

6. Stein, Lawrence H.; Matthews, Malcolm L.; and Frenk, Joel W.: "Stop - A Computer Program for Supersonic Transport Trajectory Optimization." The Boeing Co., NASA CR-793, 1967.

7. Matthews, M. L.; and Watson, D. A.: "Trajectory Optimization Program (TOP) AS 2081." Vol. 1. Analytical Development and User's Manual. No. D2-125323-1, The Boeing Co., Feb. 1968.

8. Schmidt, H. E.; Helgason, R. V.; Witherspoon, J. T.; and Geib, K. E.: "Trajectory Optimization by Method of Steepest Descent." Tech. Rep. AFFDL-TR-67-108, vol. II, U. S. Air Force, April 1968.

9. Cockayne, William G.: "X-15 Optimal Control Study." Rep. No. 2500-950002, Bell Aerosystems Co., April 1968.

10. Balakrishnan, A. V.: "On a New Computing Technique in Optimal Control and Its Application to Minimal Time Flight Profile Optimization." Presented at 2nd International Colloquium on Methods of Optimization, Novosibirsk, USSR, June 1968.

11. Taylor, Lawrence W., Jr.; Smith, Harriet J.; and Iliff, Kenneth W.: "Experience Using Balakrishnan's Epsilon Technique to Compute Optimum Flight Profiles." Paper presented at AIAA 7th Aerospace Sciences Meeting, New York, N. Y., Jan. 20-22, 1969.

12. Taylor, Lawrence W., Jr.; Smith, Harriet J.; and Iliff, Kenneth W.: "A Comparison of Minimum Time Profiles for the F-104 Using Balakrishnan's Epsilon Technique and the Energy Method." Paper presented at the Symposium on Optimization, Nice, France, June 29-July 5, 1969.

DETERMINATION OF OPTIMAL CONTROL USING IMBEDDING OF THE TERMINAL CONDITIONS

By Raymond C. Montgomery
Langley Research Center

SUMMARY

The utility of a method developed by the author to solve optimal control problems is discussed with emphasis placed on the solutions of singular optimal control problems. The method involves imbedding the optimization problem of interest in a family of optimization problems which are parameterized by their terminal conditions. The solution to the optimization problem is obtained by collapsing the terminal conditions of the family of problems onto those of the original problem, while appropriately modifying the control function to account for the changing terminal conditions. The utility of this method is illustrated by obtaining the solution to an aerial attack problem where the attacking vehicle must maneuver subject to control limitations so as to place the enemy in a certain relative position for launching a missile in the shortest possible time. This is a singular control problem. Thus, the usefulness of this method in determining the placement of singular subarc in optimal trajectories is pointed out.

DISCUSSION

It is the purpose of this paper to point out the utility of solving optimal control problems by means of imbedding the terminal conditions. This will be a brief description of the research accomplished in reference 1.

In general, optimal control theory has been developed to enable one to determine how to select the control of a process so that some quantity - called the index of performance - achieves an optimal (maximum or minimum) value during the evolution of the process. It is assumed here that the control u influences the process through a set of ordinary differential equations - $\dot{x} = f(x,u)$ - indicated in Figure 1. Here the quantity x is a list of variables which define the "state" of the system and u is a list of variables which represent the "controls" of the process. It may be required that the controls u(t) transfer the state x(t) along a path that satisfies certain prespecified initial and terminal conditions. This is illustrated schematically on the right side of Figure 1. The prespecified initial conditions and terminal conditions are indicated on the figure along with a possible trajectory x(t) which transfers x from the initial to the terminal

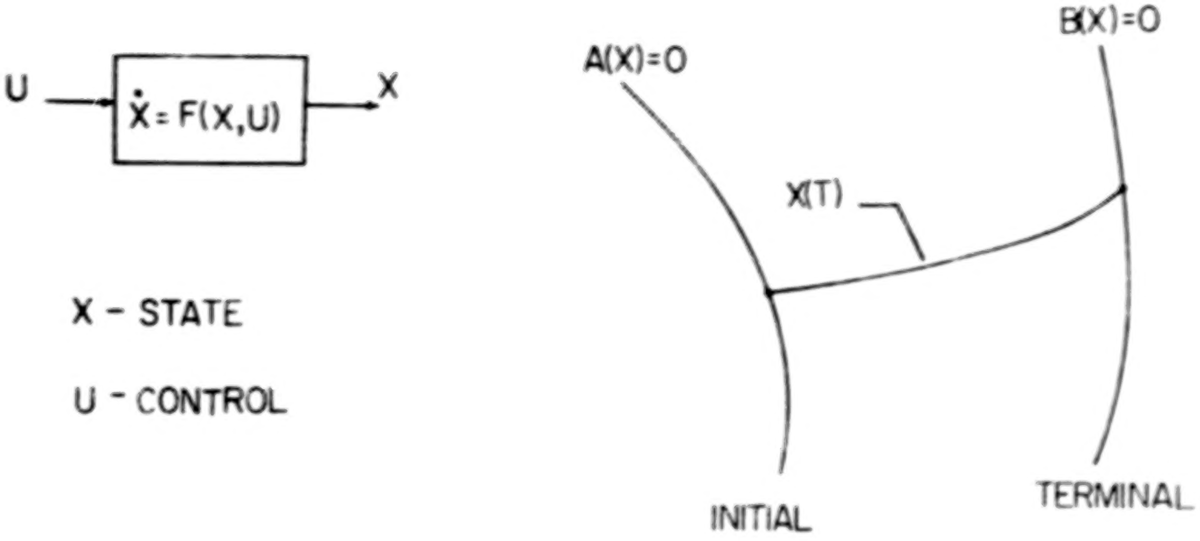


Figure 1.- Process dynamics

conditions. The problem is to make this transfer in such a way that the performance index is optimal.

There are several tools available for obtaining the solution to this problem - the calculus of variations, the maximum principle, dynamic programming, and iterative methods such as steepest descent. I will discuss yet another method for obtaining solutions to optimal control problems - imbedding. This approach has distinct advantages which will be pointed out below.

The basic idea underlying the use of imbedding to solve optimal control problems is schematically illustrated in Figure 2. The problem considered is to transfer optimally from the initial condition x^o indicated at the top right side of the figure to the terminal condition indicated by $\sigma = \sigma_f$. To solve this problem by using imbedding, the original problem is considered as one of several optimization problems which are distinguished by their terminal conditions. One member of this family should have a known solution. For a time-optimal problem one member which has a known solution might be the member corresponding to the $\sigma = 0$ terminal conditions indicated at the top right of Figure 2. Necessary conditions which allow one to pass from one member of the family to another have been derived in reference 1. These results enable one to obtain the optimal solution for one member from a knowledge of the optimal solution of another member if their terminal conditions differ only slightly. Note that the state x is a function of time t and a parameter σ which identifies the particular member of the family of terminal conditions for which the solution holds.

TABLE OF CONTENTS

1. SPACE VEHICLE CONTROL APPLICATIONS

Design Development of the Apollo Lunar Module
Kenneth J. Cox
Manned Spacecraft Center 1 1/A10

Design Development of the Apollo Command and Service Module Thrust Vector Attitude Control System
William H. Peters
Manned Spacecraft Center 63 1/E3

Orbiting Astronomical Observatory (OAO-A2) Stabilization and Control Subsystem
Thomas E. Huber
Goddard Space Flight Center 83 1/G10

Applications of the Hybrid Coordinate Method to the TOPS Autopilot
G. E. Fleischer
Jet Propulsion Laboratory 127 2/D7

Steering Laws for Double-Gimbal Control-Moment Gyros
Stephen W. Winder
Marshall Space Flight Center 157 2/F13

2. AIRCRAFT GUIDANCE, CONTROL AND HANDLING QUALITIES

Optimum Horizon Guidance Techniques for Aircraft
Heinz Erzberger and Homer Q. Lee
Ames Research Center 175 3/A8

Inertially-Aided, All-Weather Landing Systems for C/TOL.
Robert J. Pawlak
Electronics Research Center 185 3/B4

An Approach to the Determination of Aircraft Handling Qualities Using Pilot Transfer Functions
James J. Adams and Howard G. Hatch, Jr.
Langley Research Center 199 3/C4

A Frequency Domain Approach to Handling Qualities Design
William A. Wolovich
Electronics Research Center 217 3/D8

3. SYSTEM SIMULATION AND IDENTIFICATION

Optimal Regulator or Conventional? Setup Techniques for a Model Following Simulator Control System
Dwain A. Deets
Flight Research Center 237 3/F1

TABLE OF CONTENTS (CONT'D)

3. SYSTEM SIMULATION AND IDENTIFICATION (CONT'D)

Quasi-Optimal Control of a Moving-Base Simulator
Elwood C. Stewart
Ames Research Center 253 3/G3

Some Modern Control Techniques for Human Operator Modeling and Identification
Richard S. Shirley
Electronics Research Center 261 3/G11

4. ENGINE CONTROL AND ADVANCED PROPULSION TECHNIQUES

Normal Shock and Restart Controls for a Supersonic Airbreathing Propulsion System
George H. Neiner, Gary L. Cole, and Francis J. Paulovich
Lewis Research Center 299 4/D2

Optimal Control of a Supersonic Inlet to Minimize Frequency of Inlet Upstart
Bruce Lehtinen, John R. Zeller, and Lucille C. Geyser
Lewis Research Center 323 4/F1

Control System Studies for Thermionic Reactors
Richard J. Hermsen,
California State Polytechnic College
Henrik G. Gronroos
JPL, California Institute of Technology 337 4/G1

Application of Fluidics to New Control Components
Miles O. Dustin, Vernon D. Gebben, and Robert E. Wallhagen
Lewis Research Center 365 5/B7

Minimum Energy Control of a Class of Electrically Driven Vehicles
Yilmaz E. Sahinkaya
JPL, California Institute of Technology 387 5/D6

5. ADVANCED CONTROL TECHNIQUES

A New Formulation for the Epsilon Method Applied to the Minimum Time-to-Climb Problem
Lawrence W. Taylor, Jr., and Harriet J. Smith
Flight Research Center 423 5/G1

TABLE OF CONTENTS (CONT'D)

5. ADVANCED CONTROL TECHNIQUES (CONT'D)

Determination of Optimal Control Using Imbedding of the Terminal Conditions
Raymond C. Montgomery
Langley Research Center **435** 5/G13

Stability of Constant Gain Systems with Vector Feedback
George L. von Pragenau
Marshall Space Flight Center **443** 6/A10

Synthesis of Finite Settling Time Discrete Systems
James A. Gatlin
Goddard Space Flight Center **473** 6/C12

Design and Performance of Heart Assist or Artificial Heart Control Systems
John A. Webb, Jr., and Vernon D. Gebben
Lewis Research Center **495** 6/E1

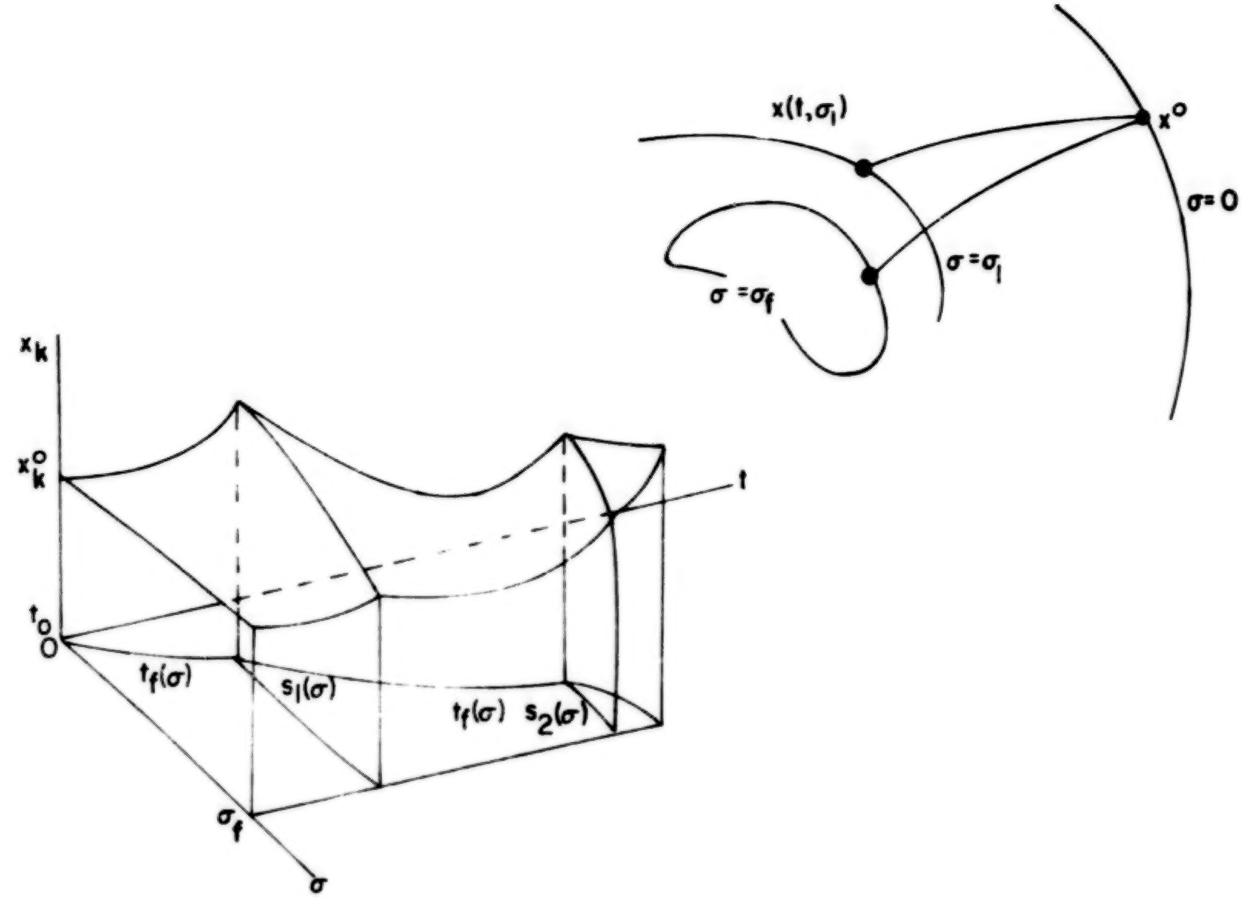


Figure 2.- Terminal imbedding concept

The general structure of the solutions is illustrated in the lower left portion of Figure 2. The typical variation of a component of the state x_k is illustrated. The variable x_k is defined over a region in t,σ plane extending from the $t = t_o$ line to the $t = t_f(\sigma)$ line as indicated. In the region of the t,σ plane between the $t = t_o$ line and the line $S_1(\sigma)$ the state is depicted as being continuous in t and σ with continuous slopes. The lines $S_1(\sigma)$ and $S_2(\sigma)$ represent the lines in t,σ space along which discontinuities normally result from discontinuous control (bang-bang control, etc.). Their location in t,σ space is not known a priori and must be determined. This is also true of the line $t_f(\sigma)$. The determination of the nature of the t,σ space is the most important step in applying imbedding to optimal control problems.

Three necessary conditions which allow determination of the nature of the t,σ space have been derived in reference 1. Two have been shown to be equivalent to the maximum principle for the imbedded set of problems. The third is a result of an assumption placed on the analysis that the field of solutions x(t,σ) has a behavior illustrated in the lower left of Figure 2; that is, assuming that the state and its slopes in t are continuous. This third condition is of use in the construction of the field of optimal controls.

Why should one resort to the use of imbedding to solve optimization problems when other methods are available? The answer to this question lies in the utility of imbedding in solving singular control problems. Generally the solution to the given optimal control problem may include two types of subarcs. This is illustrated at the top of Figure 3: regular arcs - along which the conditions of the maximum principle provide the control to be used - and singular subarcs - along which conditions of the maximum principle fail (without some extension) to define the control to be used. A procedure for determining the control to be used along a singular arc was presented by Johnson (ref. 2) using the maximum principle. The basic difficulty is however not in finding the control to be used over the singular arcs but in piecing together regular and singular arcs in such a way that the performance index is optimized and the initial and terminal conditions are satisfied.

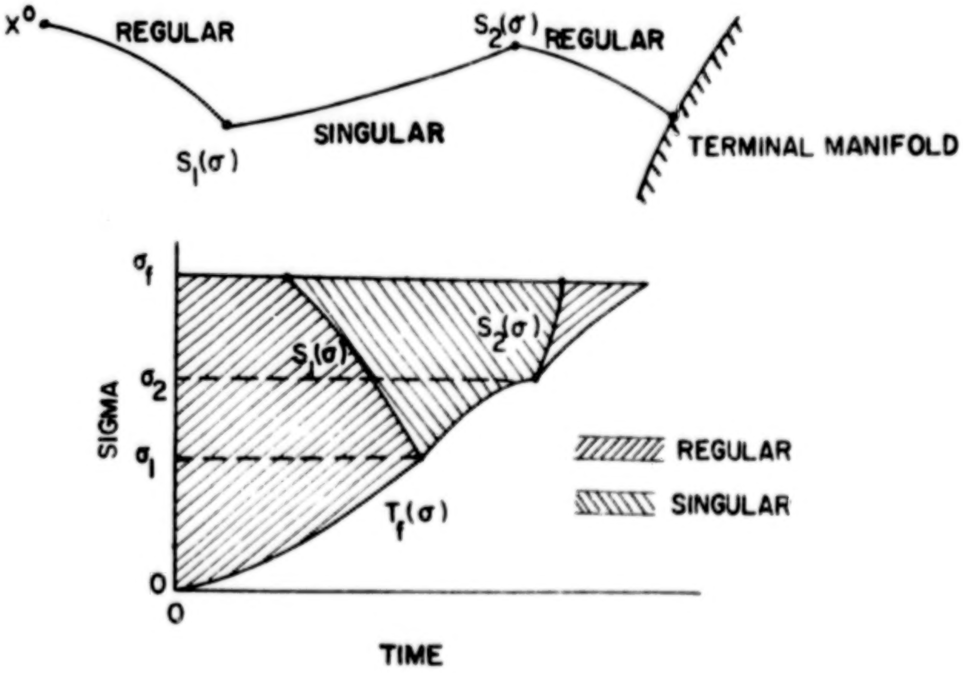


Figure 3.- Solution to singular control via imbedding

The way in which imbedding theory resolves the problem of piecing together singular and regular arcs to determine the solution to singular optimal control problems is illustrated in the lower portion of Figure 3. For a problem that the t,σ plane of

Figure 3 corresponds to the necessary conditions indicate that for values of σ between $\sigma = 0$ and $\sigma = \sigma_1$ no singular arc is contained in the solution. At $\sigma = \sigma_1$ it becomes necessary to consider the possibility of including a singular arc in the problem. It is possible at each value of σ greater than σ_1 to compute two distinct families of slopes $dS_1/d\sigma$ and $dt_f/d\sigma$ one of which assumes that it is discontinued. By direct comparison of the performance index associated with each assumption one may determine whether or not the singular arc should be continued. For this example for values of σ between σ_1 and σ_2 the singular arc was retained. At $\sigma = \sigma_2$ the singular arc was left, which created a curve $S_2(\sigma)$ separating the regular and the singular regions.

To illustrate the concepts above, the theory has been applied to an aerial attack problem whose basic features are illustrated in Figure 4. An aircraft A attacks a target T with maneuvers restricted to the horizontal plane. The mathematical model of the attacking aircraft involves a square law drag term and there is used a forward speed control. The turning acceleration of the vehicle is another control. The drag characteristics of the vehicle were taken from the F105-D airplane at 910 ft/sec airspeed and 15,000-ft altitude. The maximum speed was 1500 ft/sec and the minimum was 500 ft/sec. Turning acceleration was limited to 6 g. The target vehicle was assumed to be moving at 1000 ft/sec along the x_A axis and to be nonmaneuvering. The attacker's weapons system is a bomb which can be launched from the attacker's position.

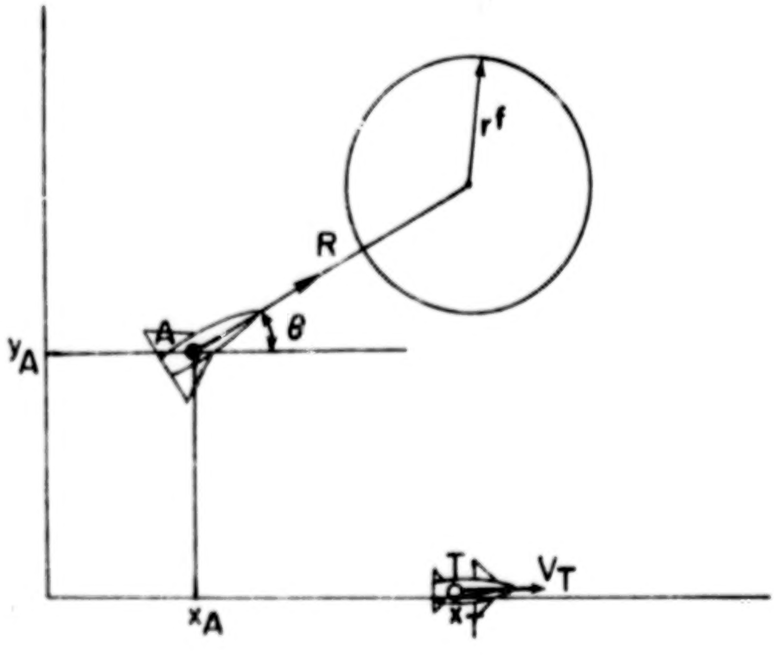


Figure 4.- Geometry of aerial attack

The bomb explodes a distance R = 3000 ft directly in front of the attacker's position with a lethal radius of r^f = 100 ft. The problem considered here was to maneuver the aircraft in such a way that the attacker places the target within the lethal radius of the bomb in the shortest possible time.

To solve this problem by using imbedding, one must first select a family of terminal conditions - one of which has a known solution and one of which is the terminal condition of the original problem statement. For the aerial attack problem the family of terminal conditions selected were terminal conditions corresponding to bombs of different radii. For $\sigma = 0$ the bomb's lethal radius was taken to be large enough to kill the enemy instantly. For $\sigma = 1$ the bomb's lethal radius was that of the original problem statement - 100 ft. Figure 5 is a graph illustrating the solutions to this problem obtained for a specific set of initial conditions. Note that the interminal positions of the problem are plotted for different values of the imbedding parameter σ. At $\sigma = 0$ the terminal conditions and initial conditions are the same, since the enemy is killed instantly. This requires a large bomb of approximately 11,000 ft lethal radius. For smaller bombs $(\sigma>0)$ the terminal conditions vary as indicated. For large bombs $(\sigma<\sigma_1)$

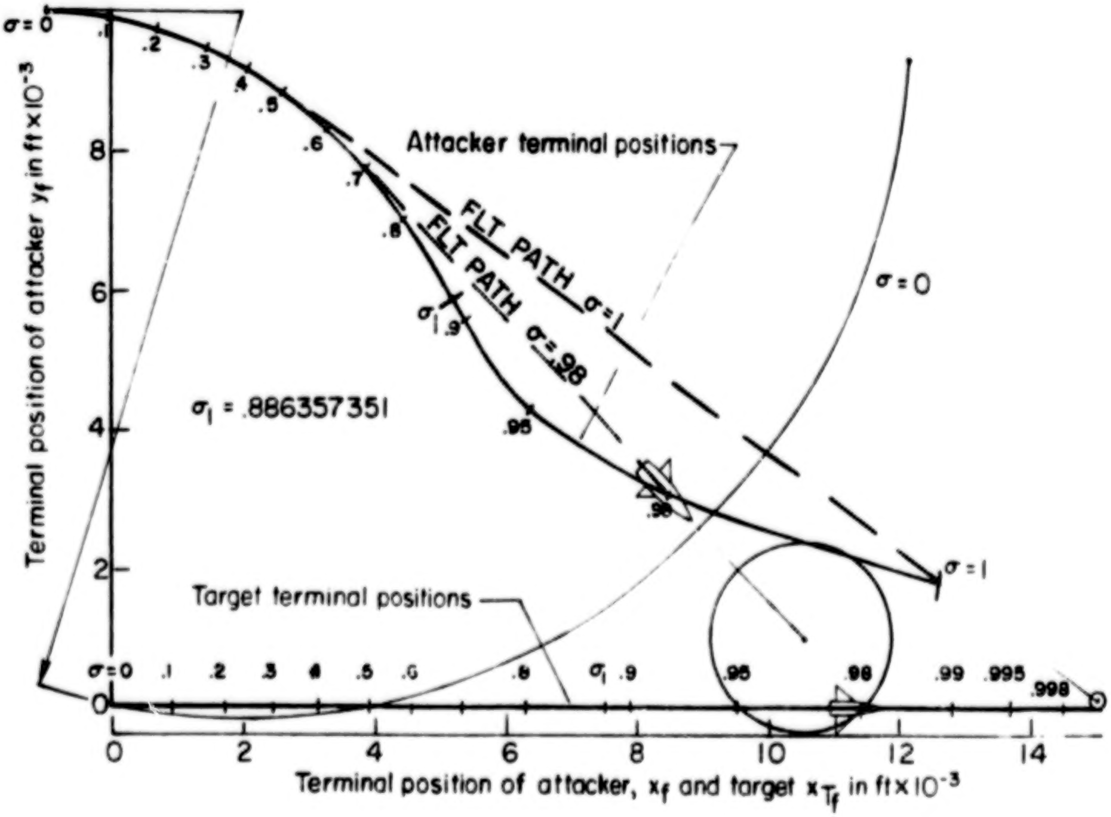


Figure 5.- The family of solutions via imbedding

only one maneuver is required - a hard right turn at full acceleration. However for smaller bombs, corresponding to values of $\sigma > \sigma_1$, two types of maneuvers are required - a hard right turn at full acceleration followed by a straight line interception course at full acceleration, which is a singular subarc. For $\sigma = .98$ the appropriate flight path is indicated on the figure. The lethal radius of the bomb is also illustrated. Note that at the instant that the attacker reaches the solid curve the target is maneuvered into the lethal radius of the bomb. The solution to the original problem ($r^f = 100$ ft) is obtained for $\sigma = 1$.

The character of the t,σ plane for this problem is illustrated in Figure 6. Note that there were two zones in the t,σ plane for this problem. The zone corresponding to $u = -1$ (corresponding to a 6 g turn) is a regular subarc zone, whereas the zone corresponding to $u = 0$ (corresponding to a straight course) is a singular subarc zone. These are separated by the line $S_1(\sigma)$, which initiates at a value of $\sigma = \sigma_1$ indicated on Figure 5. At $\sigma = 1$ the solution for turning acceleration obtained by imbedding is indicated on the right side of Figure 6. This indicates a hard right turn at 6 g for a time $S_1(1)$ and a straight line course from time $S_1(1)$ to time $t_f(1)$.

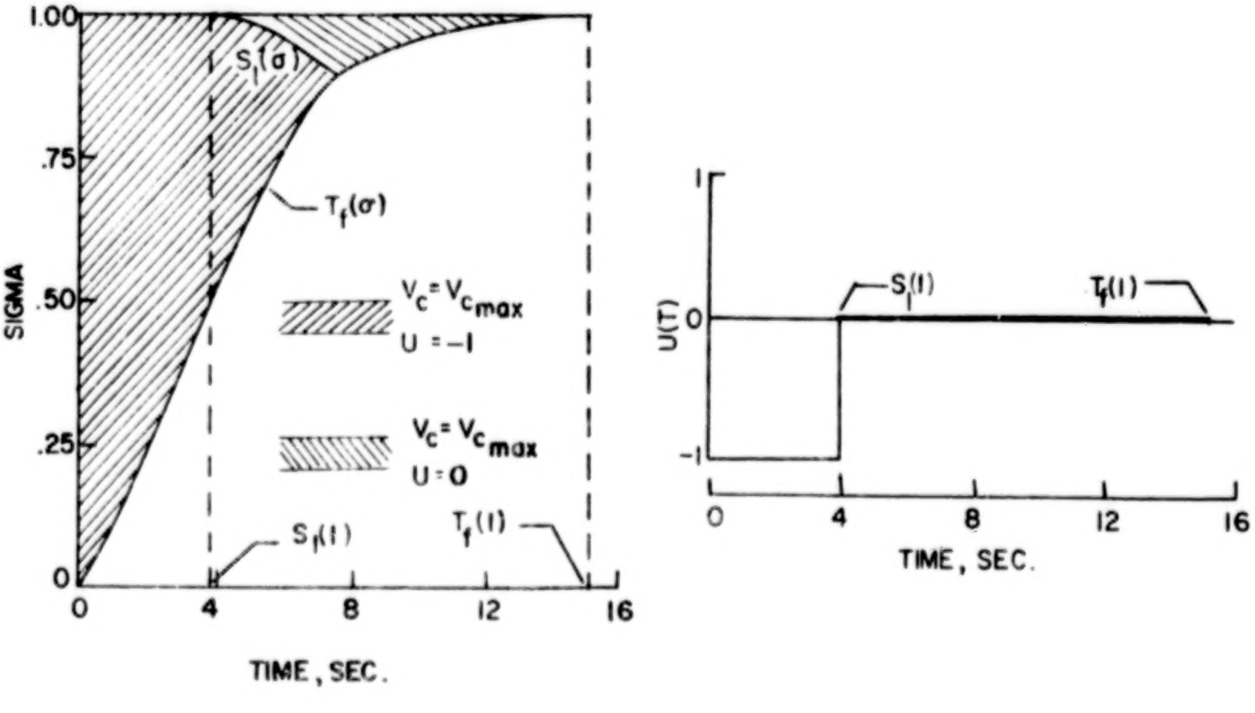


Figure 6.- Aerial attack solution in t,σ plane

In summary, necessary conditions have been derived in reference 1 that allow the application of imbedding to singular optimal control problems. Their application has been illustrated by an example of an aerial attack problem which is a singular control problem. In terms of future analytic studies regarding the use of imbedding to solve control problems, I intend to apply these concepts to the theory of differential games. Since it has been useful in studying singular optimal control problems, I feel that it has promise in differential games which are plagued with a host of singular control situations.

REFERENCES

1. Montgomery, R. C.: Trajectory Optimization by Terminal Imbedding. Ph.D. Dissertation, Virginia Polytechnic Institute, May 1969.
2. Johnson, C. D. and Gibson, J. E.: "Singular Solutions in Problems of Optimal Control." IEEE Trans on Automatic Control, January 1963, pp. 4-15.

STABILITY OF CONSTANT GAIN SYSTEMS WITH VECTOR FEEDBACK

George L. von Pragenau
Marshall Space Flight Center

SUMMARY

Multivariable feedback systems can be divided into two classes: one where all loops can be opened with a single node, and the other where this is impossible. The stability of the latter, denoted here as vector feedback, is analyzed for the class of constant gain (linear and time-invariant) systems.

The state space, the controllability, and the observability concepts are discussed in connection with the proposed stability analysis which permits drastic dimensional reductions for a vector feedback problem. Any constant gain system's stability can thus be analyzed in the frequency domain with a single Nyquist plot. The analysis considers the total system with all loops closed, a disturbance vector as input, and the feedback vector as output. All constant gain systems are shown to be decomposable into stable subsystems where the degree of the decomposition determines the dimensions. The maximum decomposition results in the state-space approach which is the limit case.

The method is demonstrated with the stability analysis of the pogo phenomenon, an oscillatory interaction between the propulsion and the structure of a space vehicle. This problem, with eigenvalues over a hundred, was drastically but rigorously reduced to a stability analysis of a 4x4 matrix which finally was evaluated with one Nyquist plot over a frequency range from 2 to 30 Hz where resonances occurred at an average density of one resonance per 1 Hz frequency interval.

INTRODUCTION

Vector feedback is defined here as a multivariable feedback loop which cannot be opened with a single node. Such systems are not necessarily built by design, but can be encountered as outgrown secondary effects as exemplified by the pogo phenomenon on our large liquid propellant rockets. There, the propulsion's main function is to propel the rocket; however, longitudinal vibrations can be sufficiently amplified by structural resonances such that propellant pressures may result in appreciable thrust oscillations which, in turn, could keep the vibrations going in a regenerative fashion. This occurred with the Atlas, Thor, Titan, and Saturn V space vehicles (refs. 1-10). The Saturn V case was also complicated by an independent motion of two thrust groups such that it became impossible to open all loops with one node. The first and second stages each have a cluster of five engines in which the center engine is one thrust group and the four outboard engines are the other. These cases can be classified as a pogo phenomenon with vector feedback (ref. 11). More details are discussed later when the proposed stability analysis technique is demonstrated.

The complexity of vector feedback often prevents rigor from penetrating the maze of applications even though we have progressed in the area of multivariable constant gain systems. Two of the major contributions are the state-space approach (refs. 12-22) and Kalman's controllability and observability (C&O) concept (ref. 23). The C&O brought the black box approach (input/output representation of a system) into the proper perspective, but it was not designed to produce stability criteria.

The state-space approach has advanced modern control theory; at least the constant gain systems seem to be understood. However, the state-space formulation has not developed into a practical design tool in most engineering areas and the frequency domain methods are still very much in use. This situation evolved from difficulties in fully understanding the problem area and the tempting fact that frequency domain methods work very efficiently. Many authors (refs. 24-28) explored the frequency domain methods for the stability analysis of multivariable systems with partial success. The results range from relatively limited applications to many confusing possibilities without producing a generally applicable and simple technique. Recent papers by Chen (refs. 29-30) give a rigorous treatment of controllable and observable multivariable feedback systems where stability is evaluated with one Nyquist plot of a determinant and the least common denominator of all minors of a transfer function matrix. The controllability and observability requirement makes the method less suitable for multivariable systems which are of relatively high order.

The method outlined here attempts to avoid the difficulties of other techniques and takes full advantage of a building block approach where the blocks are stable subsystems. The method takes into account the theoretical foundation of the state-space and the C&O concepts and gives a simple proof based on the following strategy:

1. All constant gain systems decompose into stable subsystems.
2. The total system is analyzed with all loops closed, a disturbance vector as input, and a feedback vector as output.
3. Matrix operations are available preserving the stability of subsystems, while permitting any desirable decomposition.

4. The stability of all looping connections are described by the system's determinant which is preferably evaluated with one Nyquist plot.
5. All non-controllable and non-observable subsystems are stable.

NOMENCLATURE

A,B,C	constant state space matrices
C_I, C_0	4x4 matrices relating interface forces to propellant forces. for the inboard engine, and outboard engine group, respectively
C&O	controllability and observability
D	1x2 disturbance force row vector
D_I, D_0	disturbance force at inboard and outboard engine group, respectively
D_{ik}, D_i, D_k	denominator polynomials
D_s	dashpot constant of simplified pogo model
E	thrust gain constant of simplified pogo model
f	disturbance force on simplified pogo model
G	2x1 thrust gain matrix, G_I for inboard, and G_0 for outboard
g_F, g_L	thrust gain for fuel and lox side, respectively
I	identify matrix
K	2x2 cavitation stiffness matrix, K_I for inboard, and K_0 for outboard
K_F, K_L	cavitation stiffness on fuel pump inlet and lox pump inlet, respectively
K_N	2x2 nominal cavitation stiffness matrix for eigenvalue model
K_s	cavitation stiffness on simplified pogo model
LOX	liquid oxygen oxidizer
m	structural mass of simplified pogo model

m_s	lox mass of simplified pogo model
m_i	generalized mass of pogo model for structure and propellant
N_{ik}, N_i, N_K	numerator polynomials
P	1x2 propellant force variation vector
P_{FI}, P_{F0}	propellant force variation of fuel side, on inboard, and on outboard, respectively
P_{LI}, P_{L0}	propellant force variation on lox side, on inboard, and on outboard, respectively
P_s	propellant force variation in simplified model
Q	2x2 engine flow matrix, Q_I for inboard, and Q_0 for outboard
q_{FF}, q_{FL}	flow matrix element from fuel force to fuel flow and to lox flow, respectively
q_{LF}, q_{LL}	flow matrix element from lox force to fuel flow and to lox flow, respectively
R	invertible system matrix in SMF
S	system matrix in SMF
SMF	stable matrix form
s	Laplace operator
T	transformation matrix
T_s	thrust force variation on simplified pogo model
u	input (row vector or scalar)
U	input in Laplace transform
V	similar matrix
y	state (row vector)
y(0)	initial state
Y	state in Laplace transform

ΔY_{FI} $=Y_{FI}-Y_I$	inboard fuel mode vector difference
Y_{FI}, Y_{FO}	fuel mode column vector upstream of inboard and outboard cavitation stiffness, respectively (one component per resonance)
ΔY_{FO} $=Y_{FO}-Y_0$	outboard fuel mode vector difference
Y_I, Y_0	inboard and outboard engine mode column vector, respectively (one component per resonance)
ΔY_{LI} $=Y_{LI}-Y_I$	inboard lox mode vector difference
Y_{LI}, Y_{LO}	lox mode column vector upstream of inboard and outboard cavitation stiffness, respectively (one component per resonance)
ΔY_{LO} $=Y_{LO}-Y_0$	outboard lox mode vector difference
z	outboard (row vector or scalar)
Z	output in Laplace transform
α_i	integrated impulse value
λ_i	eigenvalues
Λ	diagonal matrix of eigenvalues
Ω^{-1}	diagonalized matrix of pogo model for structure and propellant
ω_i	resonance of pogo model in radians
ζ_i	critical damping of pogo resonances

STATE SPACE

The state-space representation is recalled here to illustrate certain points and steps taken in the following sections. The dynamic behavior of any constant gain system can be described by a first-order vector differential equation with constant coefficient matrices. This paper uses variables in row vector form which transposes matrices of other common notations. The row vector format was selected because it best resembles mappings and flow diagrams.

$$\dot{y} = yA + uB \tag{1}$$

$$z = yC. \tag{2}$$

The solution of Eq. (1) is:

$$y(0) \exp [At] + \int_{o}^{t} uB \exp [A(t - \tau)] \, d\tau = y. \tag{3}$$

The state-space formulation [Eqs. (1) and (2)] assumes that there exists a state vector y which completely describes the interior state of a system, an assertion which is often not easily verified unless the internal structure of the system is known. The input u and the output z are, in general, only partially related to the state vector y; consequently, input and output measurements do not always represent the state y or the total system. This question is directly related to the controllability and observability concept.

The matrices A, B, and C are constant; the exponential matrix functions $\exp[At]$ and $\exp [A(t-\tau)]$ are time-dependent and are denoted as state transition matrices [Eq. (3)]. The stability depends on the transition matrix which must be evaluated, e.g., by obtaining the eigenvalues of the matrix $A(|A-\lambda I| = 0)$. The matrix A can then be presented by a similarity transformation TJT^{-1} which also transforms the transition matrix $\exp [At] = T\{\exp[Jt]\}T^{-1}$ where J is the Jordan canonical form. The function $\exp [Jt]$ is now readily reduced to a matrix with exponential functions $\exp (\lambda_i t)$ (and/or their derivatives) as elements. Stability prevails if all eigenvalues λ_i have negative real parts, meaning that the state y approaches zero from any initial state $y(0)$ when time grows unlimited and when the input u is zero [Eq. (3)].

Thus, the state-space method led directly to a stability criterion without leaving the time domain; and we wonder if an earlier adoption centuries ago would not have severely curtailed the popularity of differential operators. However, eigenvalues are relatively difficult to obtain for large order systems, especially when we deal with the non-conservative type; therefore, it appears worthwhile to explore other avenues such as the frequency domain approach.

We now turn to the Laplace transform of Eqs. (2) and (3):

$$y(0)[sI - A]^{-1} + UB[sI - A]^{-1} = Y \tag{4}$$

$$YC = Z. \tag{5}$$

The transformed variables are U, Y, and Z; input U and output Z are vectors or scalars and state Y is a vector. The initial state y(0) and the input U are mapped by the matrix $[sI-A]^{-1}$ into the state Y which again is mapped into the output Z (Figures 1 and 2).

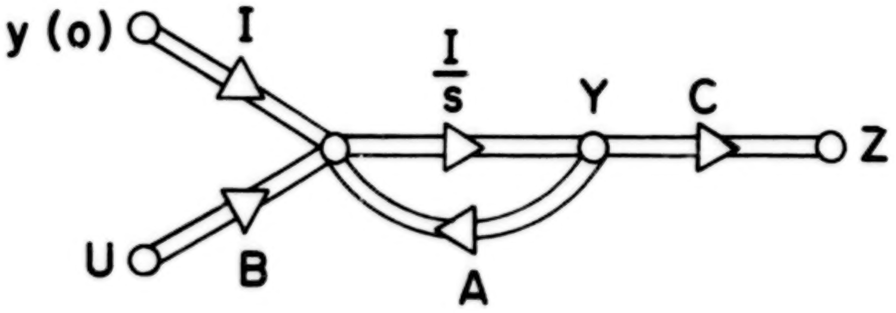


Figure 1.- State-space flow diagram with vector feedback by Y; U and Z are the input and the output vectors, respectively.

Figure 1 results from the Laplace transform of Eq. (1). Note that the state-space case is a typical vector feedback situation. Another flow diagram results from Eqs. (4) and (5) where the loop is eliminated by a matrix inversion.

The simplicity of the input-output concept is appealing but this should not detract our attention from the controllability and observability concept which requires that the state is also considered. To investigate stability, the matrix inverse is arranged in the form of the classical adjoint.

$$[sI - A]^{-1} = \operatorname{adj}[sI - A]/|sI - A|. \tag{6}$$

All zeros are distributed among the elements of the adjoint, while all poles are collected in the determinant. Pole-zero cancellations are possible.

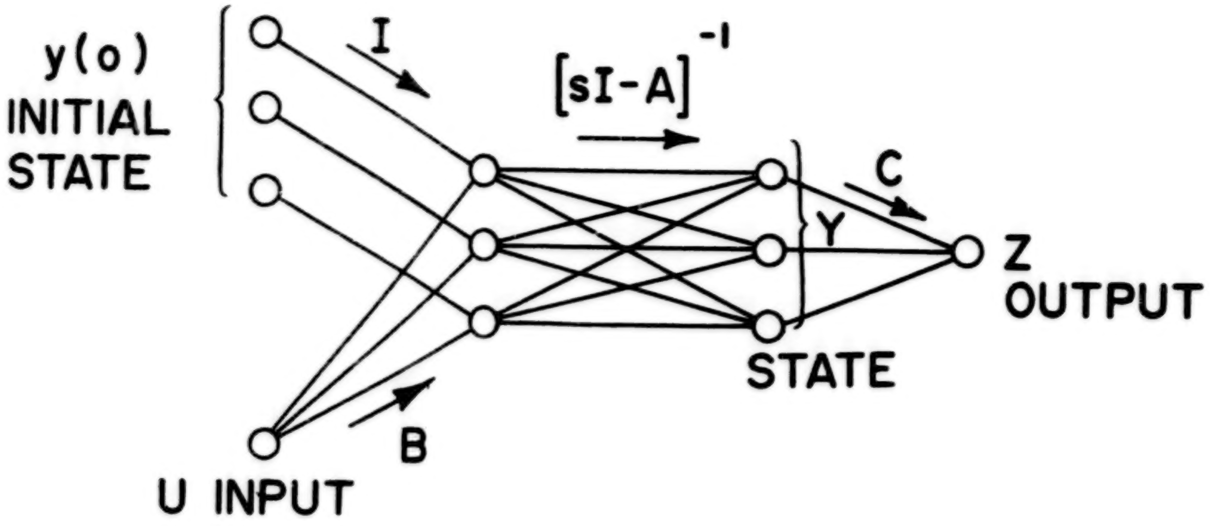


Figure 2.- State-space flow diagram for a third-order system (input and output are scalars here).

If we ignore the time domain's result and wish to discuss stability in the frequency domain only, then we become immediately involved in the details of pole-zero cancellations. First, we state that stability is certain if all roots of the determinant are stable, but the necessity is still in question. First note that the order of the determinant is at least one order higher than the order of the adjoint, thus preventing a complete cancellation. Second, the elements vary and cancellations will change and different poles will become effective for each element.

Great precision is obtained by analyzing the row vector $y(0)\text{adj}[sI-A]$ where A is assumed to represent a system in the phase space. The matrix A is then a companion matrix (Frobenius matrix) which results in an adjoint where each element within any column has roots that differ from element to element. Any initial state $y(0)$ combines these elements in a linear fashion and thus produces any desirable zeros for the components of the vector $y(0)\text{adj}[sI-A]$. Consequently, cancellations become a special case that require certain initial states. But stability should hold for any initial state (or input) and we conclude that the roots of the determinant $|sI-A|$ are not only sufficient but also necessary for representing the poles of the total system.

Both representations of the state-space approach were presented in preparation for the steps shown in the following sections; e.g., the interpretation of the matrix $[sI-A]^{-1}$ is similar except that the elements will have poles added.

CONTROLLABILITY AND OBSERVABILITY

The controllability and observability concept (C&O) poses a well-defined question about obtaining insight into a system from its outside. The controllability permits us to stimulate all states of a system through its inputs and the observability guarantees that all states can be obtained from output data. More details on definitions can be found in the literature (refs. 13,17,18).

If we assume that matrix A of Eq. (1) is a diagonal matrix Λ, then the C&O results in the simple mathematical statement that the columns of B and the rows of C must match with the dimensions of the state vector y. The C&O criterion for non-diagonal matrices A is more complicated, but the simple case is sufficient to illustrate the role that the C&O plays in relation to the stability problem.

For example, it sounds very contradictory to talk about the "controllability and observability" of an unstable system. However, the C&O concept permits such situations because it does not restrict the Λ matrix. This matrix has the eignevalues at its diagonal, and unstable eigenvalues are not ruled out. The C&O of unstable systems means only that the unstable system can be kept under "control" with a suitable input which, in turn, requires an output that "observes" all states such that any runaway can be detected and properly counteracted. This, however, is not a stability criterion; it is a classification criterion that selects from the class of unstable systems the candidate systems which can be stabilized by adding an external feedback device. Optimum control enthusiasts occasionally seem to overlook this fact.

If we consider systems solely composed of stable subsystems, then the C&O concept is of no consequence below the subsystem's level. Gilbert (ref. 31) has shown the partitioning of a system into four subsystems which are:

1. Uncontrollable and unobservable (isolated);
2. Uncontrollable but observable (partially connected);
3. Controllable but unobservable (partially connected); and
4. Controllable and observable.

If those four subsystems are stable, then the total system will be stable because isolation (1) and partial connections (2 and 3) obviously will not affect the overall stability. In cases where the stability of the four subsystems is not clear, further decomposition into stable subsystems could be performed for a stability analysis. The introduction of stability at the subsystem level permits the elimination of the non-C&O parts, thus reducing dimensions and concentrating on stability-sensitive parts.

STABLE MATRICES

The assumption that a system consists of stable subsystems is not only practical from the engineering standpoint but also proper for general applications. Most subsystems, such as elastic structures, electrical filters, or hydraulic components, are stable. If the stability status should be unclear or if a subsystem should be unstable because of some regenerative features, then a further decomposition will always lead to stable subsystems or at the worst to "drift components," which are defined here as components with poles at the origin as it occurs when some of the eigenvalues of the matrix A are zero.

The state-space approach is reached when the decomposition is carried all the way to the elements of a system where redundant elements are lumped together. However, the state-space method should be considered only as a last resort; the dimensions could be prohibitive. In other words, it is more economical to use stable subsystems as building blocks for a stability analysis whenever possible. The following strategy is recommended:

1. Divide the total system into stable subsystems;
2. Discard the non-C&O part at each stable subsystem level;
3. Keep the order of the subsystems as high as stability permits;
4. Represent the subsystem at the scalar level by stable transfer functions; and
5. Collect these transfer functions in matrix formats.

Having followed this outline, we obtain vector input-output relations (Figure 3):

$$[U_1U_2\ldots U_n]\begin{bmatrix} N_{11}/D_{11}\,N_{12}/D_{12} & \cdots & N_{1m}/D_{1m} \\ N_{21}/D_{21}\,N_{22}/D_{22} & \cdots & N_{2m}/D_{2m} \\ \vdots\qquad\quad\vdots\;\;\vdots & & \vdots \\ N_{n1}/D_{n1}\,N_{n2}N_{n2} & \cdots & N_{nm}/D_{nm} \end{bmatrix} = [Z_1Z_2\ldots Z_m] \qquad (7)$$

which can be reduced to $US = Z$. (8)

The initial conditions in Eq. (7) are neglected because their influence diminishes with time since we have only stable subsystems. The matrix S is defined as stable if all the roots of the denominators D_{ik} are stable. The matrix is therefore unstable if at least one D_{ik} is unstable.

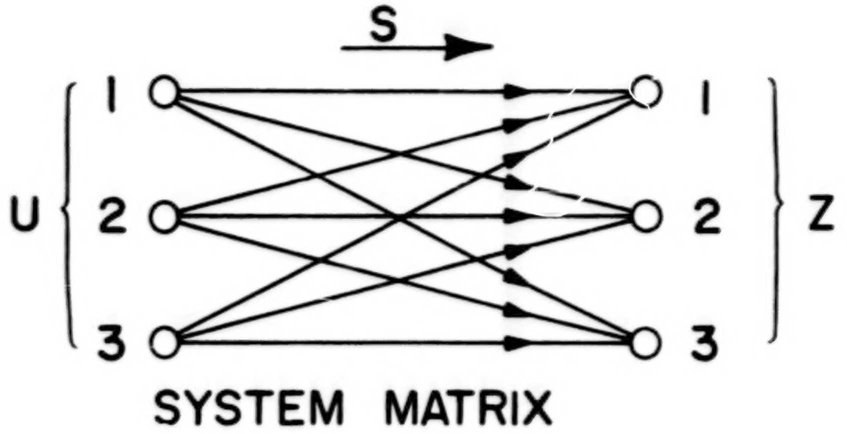


Figure 3.- Flow diagram of a three-dimensional system S mapping the input U into the output Z.

If we take as input an array of impulses $U = [\alpha, \alpha_2, ..\alpha_k ...]$, then the components of the output vector Z_i become linear combinations of the columns of S:

$$Z_i = \sum_k \alpha_k N_{ki}/D_{ki} = N_i/D_i \tag{9}$$

The factors of the common denominator $D_i = \prod_k D_{ki}$ generally will not cancel with the numerator since the elements N_{ki}/D_{ki} are all different and are linearly combined by the α_i's in an arbitrary way. Therefore, the stability within the system S is completely described by denominators of each element.

STABLE OPERATIONS

The decomposition into stable subsystems is based on several matrix operations which preserve the property of stable matrices. The operations are simple algebraic relations but are discussed because of their importance. If Σ denotes the set of stable systems and $\{S, S_1, S_2\} \subset \Sigma$, then the following theorems apply:

1. Addition $S_1 + S_2 \in \Sigma$
2. Matrix product $S \cdot S_2 \in \Sigma$
3. Adjoint $\mathrm{adj}S \in \Sigma$
4. Determinant $|S| \in \Sigma$.

Only additions or multiplications, or both, are applied to form new matrix elements. The common denominator of any new element is the product of stable denominators of previous elements and therefore the new elements must be stable again.

The roots of the newly formed numerators are of no consequence here; we observe only that when stable they could cancel some of the roots of the denominator (a very unlikely coincidence). When unstable they could not cancel any roots of the denominator because the denominator is stable. In any case, the numerators do not contribute to the stability property of the stable operations.

MATRIX INVERSION

To obtain analytical maneuverability, we must add one more operation: the matrix inversion. This operation does not belong to the category given in the previous section because it could produce instability. The inverse is presented similar to the state-space approach by the adjoint:

$$S^{-1} = \mathrm{adj}S/|S| \tag{10}$$

The adjoint and the determinant are stable, as shown previously; however, the stability of the inverse of the determinant must be investigated, e.g., by a Nyquist plot about the origin. Since the poles of $|S|$ are stable, the encirclements equal the number of unstable zeros (refs. 32,33). Stable zeros of $|S|$ ensure the stability of S^{-1}. This is definitely a sufficient condition, but the necessity is still in question because of possible cancellations between the zeros of adjS and $|S|$.

However, the cancellations can be prevented by rearranging the system S with a similarity transformation and by considering an impulse array as input. First a system is described where input U' is mapped by the system S^{-1} into an output Z':

$$U'S^{-1} = Z'. \tag{11}$$

Then the input and the output are equally transformed by a constant non-singular matrix T:

$$UT = U', \; ZT = Z' \tag{12}$$

Substituting this into Eq. (11) yields:

$$UTS^{-1}T^{-1} = Z \tag{13}$$

and we find a similar system V

$$V = TST^{-1},\ V^{-1} = TS^{-1}T^{-1} \tag{14}$$

$$V^{-1} = \mathrm{adj}V/|V| = T\{\mathrm{adj}S\}T^{-1}/|S|. \tag{15}$$

Note that the determinants are alike ($|V| = |S|$ analogous to a similarity transformation); however, the adjoints are different. The transformation T can be used, at least, to fill out one column of adjV with elements of different zeros if adjS does not already satisfy this condition.

Let us assume one column of adjV having obtained the structure

$$\mathrm{adj}V_k = [N_{1k}/D_{1k},\ D_{2k}/D_{2k},\ \dots\ N_{nk}/D_{nk}]^T. \tag{16}$$

Then, we take an impulse array as input

$$U = [\alpha_1,\ \alpha_2,\ \dots\ \alpha_n] \tag{17}$$

and form the product

$$U\mathrm{adj}V_k = \sum_i \alpha_i\ N_{ik}/D_{ik} = N_k/D_k. \tag{18}$$

The k'th output component is now

$$Z_k = N_k/\{D_k|S|\}. \tag{19}$$

The component Z_k indicates whether a cancellation between the zeros of N_k and $|S|$ is possible. First, we discard D_k which is the common denominator of Eq. (18). All the roots of D_k must be stable as initially assumed. Also, the poles of $|S|$ must be stable (see "Stable Operations"). But the roots of N_k can be freely chosen by Eq. (18) where different polynomials $N_{ik}\{\prod_l D_{lk}\}_{\partial i}$ (∂i means i excluded) are linearly combined by any impulse amplitude α_i selected. Therefore, it is always possible to find some U [Eq. (17)] which prevents cancellation of the roots of N_k and the zeros of $|S|$.

Since stability should not depend on a particular distribution of test impulses at the input, we must conclude that the zeros of $|S|$ are not only sufficient but also necessary for a stability criterion of a system S^{-1} where the subsystems or elements are stable.

The necessary and sufficient stability conditions for matrix inversions are $S \in \Sigma$ and $\mathrm{Re}\{s: |S(s)| = 0\} < 0$.

STABLE MATRIX FORM

It is always possible to arrange a constant gain system into a stable matrix form (SMF). The SMF has matrices with stable transfer function elements permitting or resulting from stable operations as outlined previously. A crucial operation is the matrix inversion which represents the stability effect of feedback and of series (or parallel) connections, depending on the involved system components which can be either unilateral or bilateral. Necessary and sufficient conditions are: (1) the decomposition of constant gain systems into SMF, and (2) the inclusion of any feedback parts into the stability analysis by a matrix inversion.

Before we delve further into details, some remarks on the state-space approach are in order. The state space requires a complete knowledge of the system. Measurements of the input and the output will only reconstruct the C&O part. Sooner or later we end up with the assumption that our knowledge is complete. If our assumptions are wrong, then we will be in for some surprises; e.g., the system burns up because an unknown system part was unstable. If we have a second chance, then this new experience will guide us in improving our analysis; however, thoroughness in analysis is recommended because it is usually less costly.

For the proposed SMF method, a complete knowledge of the system is also necessary. The SMF results in fewer dimensions than the state-space. Actually the reductions are tremendous but there is a trade-in: for the dimensional reduction, we must accept complicated transfer functions as matrix elements. The part of the system requiring stability investigation must be represented by the SMF, and the consequences of wrong assumptions are the same as for the state space. The SMF lends itself to stabilizing a system in parts (ref. 34).

Any constant gain system with stable subsystems can be Laplace-transformed and presented in the following stable matrix form:

$$US = ZR \tag{20}$$

The initial conditions are neglected because stable subsystems will eventually dampen out any transience of the initial state. The row vectors U and Z are the assumed input and output, respectively; S is one system part and R is the other. If the equations are properly arranged, then matrix R can be inverted and Z can be expressed explicitly:

$$USadjR/|R| = Z \tag{21}$$

Stability can be found by evaluating the zeros of the determinant R. However, it is possible that the inputs and outputs are unintentionally mixed which could result in a singular matrix R. Then we must convert the system by partitioning the variables and matrices according to input (index 1) and output (index 2).

$$\begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} Z_1 & Z_2 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \tag{22}$$

From this, a non-singular matrix is obtained on the right.

$$\begin{bmatrix} U_1 & Z_1 \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ -R_{11} & -R_{12} \end{bmatrix} = \begin{bmatrix} U_2 & Z_2 \end{bmatrix} \begin{bmatrix} -S_{21} & -S_{22} \\ R_{21} & R_{22} \end{bmatrix} \tag{23}$$

However, having succeeded in producing a non-singular matrix R on the right does not mean that U is an input. An input must fit the physical environment; the causality question is involved here and must be observed especially when we consider disturbance vectors as an input. Therefore, some interpretation is required before we accept a vector as an input. Similar considerations apply to the output. For example, if we consider the stability of a body, then it is understood that an external force is the input (disturbance) and that the output of interest is the displacement. Often an exchange of inputs and outputs stabilizes unstable systems and vice versa. This situation can also be interpreted such that the choice of variables implies certain boundary conditions.

POGO PHENOMENON

The stability analysis is now applied to the pogo phenomenon which was experienced within the past 10 years on practically all large liquid propellant rockets: the Atlas, Thor, Titan, and finally the Saturn V. The rockets usually oscillated with their ends moving against each other, like a youngster does on a pogo stick. The comparison led to the term "pogo effect." Falling in this category are, for example, the pressure regulator feedback on the Atlas and the propulsion-to-structure feedback on the Thor, Titan, first stage of the Saturn V, and more recently a so-called mini pogo on the second stage of the Saturn V involving only the center engine area.

The analysis demonstrated permits a survey of the stability status in a wide frequency range, but is restricted to a constant gain model which cannot reproduce the limit cycle effect observed in flight. Most likely, the limit cycle was caused by a tuning/detuning of time variable resonances (predominant on the first flight stage, S-IC), a non-linear loop gain which decreases at higher amplitudes (well-pronounced on the second stage, S-II), or possibly both. In most cases, the pogo oscillations had a football-shaped envelope, building up slowly like a slightly unstable linear system before damping out.

Before plunging into the complexity of the problem, it is very instructive to study the simple model shown in Figure 4. This model consists of the essential ingredients of the pogo phenomenon: propellant mass, cavitation stiffness at the engine's pump inlet, orifice effect of pump inlet, vehicle mass, and the thrust's sensitivity to propellant pressure at the pump's inlet.

The pogo loop can best be described through the involved components, starting with a small thrust change T_s which forces the vehicle and propellant masses to accelerate. The accelerated propellant mass, here the liquid oxygen in the tank, exerts a pressure force onto the cavitation stiffness K_s which transmits it to the pump's inlet. This pressure affects the propellant's flow which feeds the combustion in the rocket motor; thus, the thrust changes with pressure. Then the whole process starts again in a feedback fashion.

The propulsion system's sensitivity is simply described by a constant gain E which relates the propellant force P_s and thrust T_s by $P_sE = T_s$. Actually, all variables represent small variations about a quiescent point. Also an external disturbance force f at the thrust point is introduced. A closed loop equation results:

$$P_s = \frac{f}{1 + \frac{m}{m_s} - E + s\frac{m}{D_s} + s^2 \frac{m}{K_s}} . \tag{24}$$

This equation gives the response of the propellant force P_s to a disturbance f. It results that the constant $1 + \frac{m}{m_s} - E$ must be positive for stability; i.e., gain E is limited by the stability criterion

$$E < 1 + \frac{m}{m_s} . \tag{25}$$

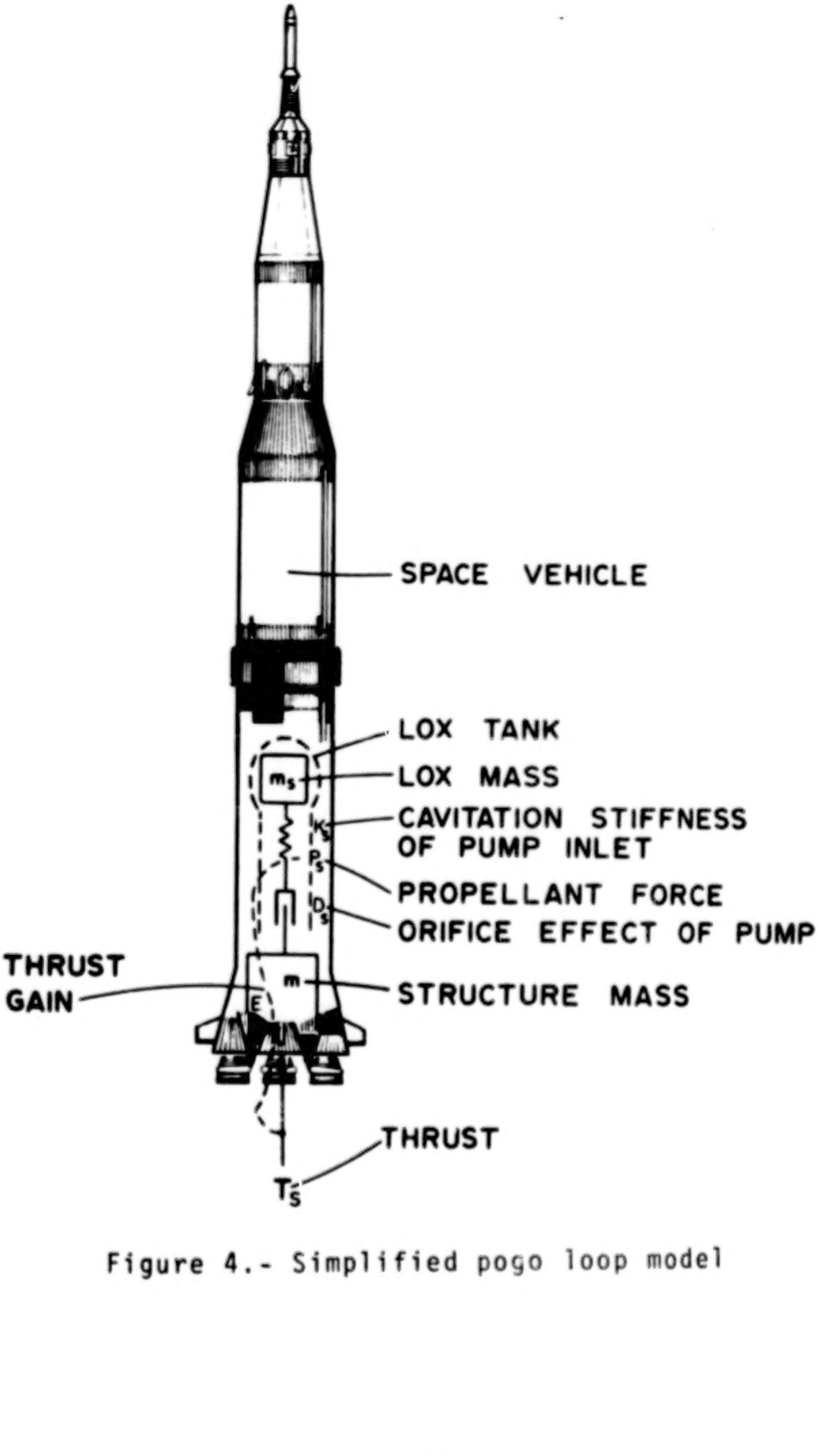

SPACE VEHICLE
LOX TANK
m_s
LOX MASS
K_s
CAVITATION STIFFNESS OF PUMP INLET
P_s
PROPELLANT FORCE
D_s
ORIFICE EFFECT OF PUMP
THRUST GAIN
E
m
STRUCTURE MASS
THRUST
T_s

Figure 4.- Simplified pogo loop model

The engine's gain should not be too high; in fact, it should be less than one if the mass ratio of vehicle structure to liquid oxygen is assumed to be negligible. All Saturn V vehicles have an E gain above one, leading to a potential pogo stability problem. However, it must be emphasized that this analysis is overly simplified and is given for illustration purposes only. A reduction of the gain E below one would be ideal, but for practical reasons stability was attained by placing a helium accumulator near the lox pump inlets where it attenuated propellant pressure oscillations (ref. 5). The pogo oscillations were successfully eliminated for the first flight stage of the Saturn V vehicles beginning with the first moon flight, the AS-503.

Actual derivation details of the pogo model can be found in a recent publication (ref. 11); this discussion is restricted primarily to the vector feedback problem.

POGO WITH VECTOR FEEDBACK

The pogo loops of the Saturn V's first and second stages (S-IC and S-II) are a case of multivariable feedback in which the feedback variable is not a scalar but a vector. The feedback is given by a matrix with plenty of crosscoupling as shown in Figure 5. The elements of the matrix are transfer functions which are more complicated than the single s operator in the diagonal elements of a state transition matrix [Eq. (6)].

The linearity and the time invariance of the used model permitted the inclusion of many resonances (30) which are relatively dense in the pogo case (1 resonance/1 Hz). The approach appears to be well justified to predict stability because the unstable buildup of actual pogo oscillations resemble oscillations of slightly unstable linear systems and because the vehicle resonances change rather slowly during flight. The non-linearities are assumed monotonic without any abrupt changes as in the case of bang-bang control systems.

Lack of knowledge in treating matrix feedback cases has often led to simplifications which caused doubt about the validity of the analysis. For example, Nyquist plots are sometimes obtained by opening one loop only while other loops remain closed. The interpretation of the stability margins then becomes obscured, mainly because of the unpredictable influence of the closed loops in the "open loop system." The stability status of the closed loops must then be analyzed by additional Nyquist plots or root-finding routines, thus producing several stability margins which cannot be combined into a single term. Variation of parameters can help to find the stability limit, but this method still does not preclude a possible pole-zero cancellation if only one open-loop plot is evaluated.

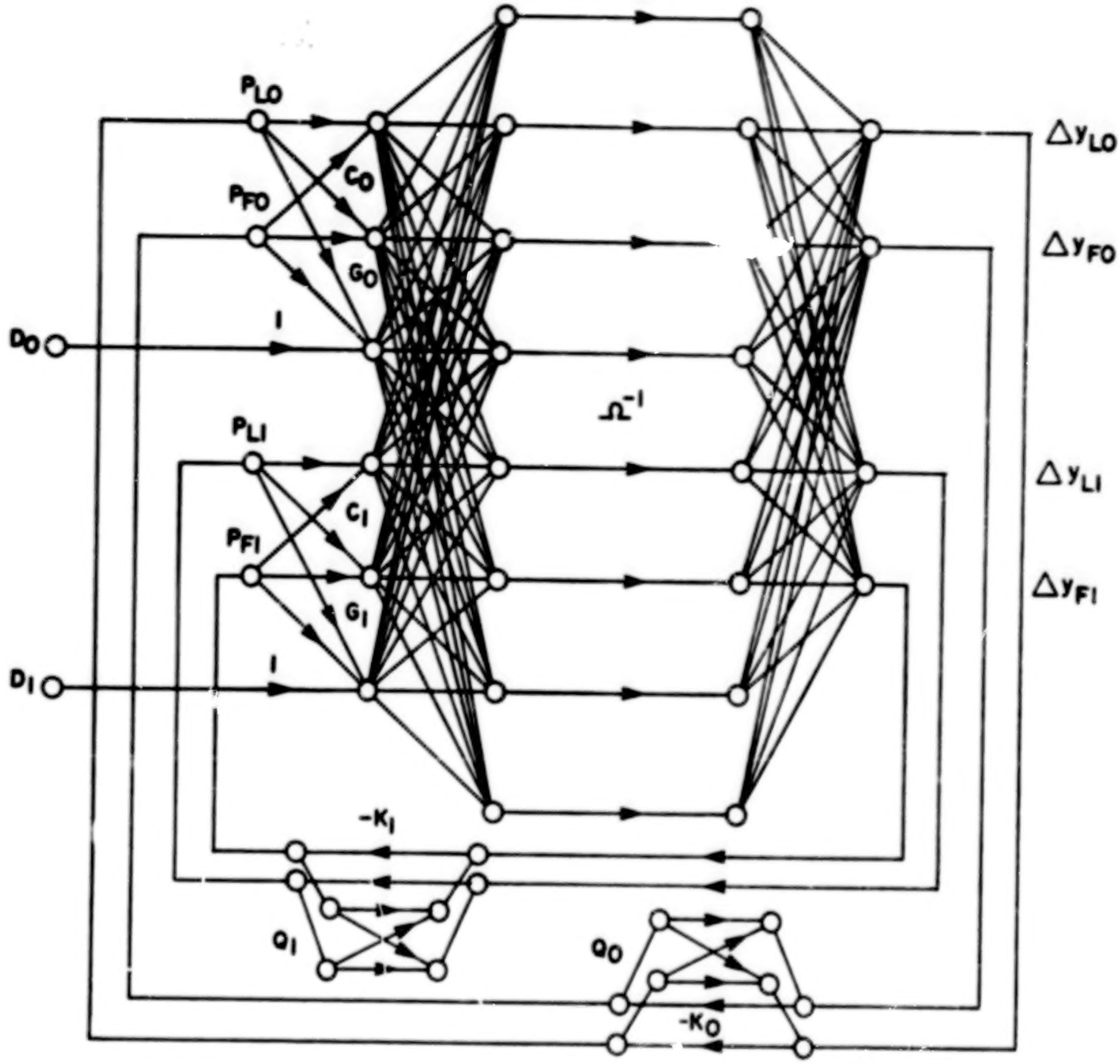

P_{LO}
P_{FO}
C_O
G_O
D_O
I
P_{LI}
P_{FI}
C_I
G_I
D_I
I
Ω^{-1}
Δy_{LO}
Δy_{FO}
Δy_{LI}
Δy_{FI}
$-K_I$
Q_I
Q_O
$-K_O$

Figure 5.- Pogo loop flow diagram with eight vehicle resonances.

The stable matrix form (SMF) is used here to analyze the pogo stability of the Saturn V's first stage S-IC. We have essentially two thrust groups moving relatively independent because of sufficient flexibility within the thrust frame. One thrust group is the inboard engine; the other is the sum of four outboard engines. This is a typical case of vector feedback as described in SMF by the equation:

$$DS = PR. \tag{26}$$

The closed-loop concept is applied with the disturbance vector D as input and the feedback vector P as output. S and R are matrices according to the rules of the SMF. The input vector is $D = [D_0 D_I]$ where D_0 is an external force on all outboard engines together, and D_I is an external force on the inboard engine. These engine points do not coincide with vibration nodes and therefore permit an excitation of all longitudinal vibration modes. External force inputs are proper because boundary conditions are thus preserved. The output vector is $P = [P_{LO} P_{FO} P_{LI} P_{FI}]$ where the indexed P's are propellant pressure forces for lox with index L, fuel with index F, outboard with index O, and inboard with index I. R and S are stable matrices as defined below:

$$R = \begin{bmatrix} Q_0+K_0^{-1} & & 0 & 0 \\ & & 0 & 0 \\ 0 & 0 & Q_I+K_I^{-1} & \\ 0 & 0 & & \end{bmatrix} + \begin{bmatrix} C_0 & G_0 & 00 & 0 \\ & & 00 & 0 \\ 00 & 0 & C_I & G_I \\ 00 & 0 & & \end{bmatrix} \begin{bmatrix} \Delta Y'_{L0} \\ \Delta Y'_{F0} \\ Y'_0 \\ \Delta Y'_{LI} \\ \Delta Y'_{FI} \\ Y'_I \end{bmatrix} \Omega^{-1} \left[\Delta Y_{L0} \Delta Y_{F0} \Delta Y_{LI} \Delta Y_{FI}\right] \tag{27}$$

$$S = - \begin{bmatrix} Y'_0 \\ Y'_I \end{bmatrix} \Omega^{-1} \left[\Delta Y_{L0} \Delta Y_{F0} \Delta Y_{LI} \Delta Y_{FI}\right]. \tag{28}$$

The Q's are 2x2 matrices of the rocket engine's pump flow versus pressure forces. The index rules are the same as for the pressure force vector P. The elements q are stable transfer functions, thus satisfying the SMF requirement. The q's are maximally second/fourth order quotients:

$$Q = \frac{1}{s} \begin{bmatrix} q_{LL} & q_{LF} \\ q_{FL} & q_{FF} \end{bmatrix} . \tag{29}$$

The K's are 2x2 matrices of the cavitation stiffnesses:

$$K = \begin{bmatrix} K_L & 0 \\ 0 & K_F \end{bmatrix} . \tag{30}$$

A nominal cavitation matrix is defined as K_N:

$$K_N = \begin{bmatrix} K_{LN} & 0 \\ 0 & K_{FN} \end{bmatrix} . \tag{31}$$

The C's are 2x2 matrices coupling pressure forces to structural forces:

$$C = I - (Q + K^{-1})K_N . \tag{32}$$

The G's are 2x1 matrices relating propellant pressure forces to thrust. The elements g are stable transfer functions which are maximally second/third order quotients:

$$G = \begin{bmatrix} g_L \\ g_F \end{bmatrix} . \tag{33}$$

The ΔY's are longitudinal mode shapes with the same index rules as for the pressure vector P; e.g., LO means outboard lox fluid displacement versus the outboard engine, O refers to the outboard engine only. Each ΔY is actually a column vector, one component per longitudinal resonance; $\Delta Y'$ is the transpose. The same holds for the Y's.

The Ω^{-1} is a diagonal matrix of stable second-order systems which represents 30 longitudinal resonances with the generalized masses m_i, criterial damping ζ_i, and circular frequency ω_i:

$$\begin{bmatrix} \frac{1}{s^2 m_1} & 0 & 0 & \cdot \\ 0 & \frac{1}{(s^2 + s2\zeta_2\omega_2 + \omega_2^2)m_2} & 0 & \cdot \\ 0 & 0 & \frac{1}{(s^2 + s2\zeta_3\omega_3 + \omega_3^2)m_3} & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \tag{34}$$

In reviewing our Eqs. (29)-(34), we find building block matrices with only stable elements. These matrices are then combined by stable operations in Eqs. (27) and (28), thus representing a true SMF structure. The final step in the inversion of R:

$$D\ S\ \mathrm{adj}R/|R| = P \tag{35}$$

Note that R is a 4x4 matrix and that the matrices of Eq. (27) are from, left to right: 4x4, 4x6, 6x30, 30x30, and 30x4.

Now a Nyquist plot about the origin from the determinant $|R|$ is computed.* The number of the clockwise encirclements equals the number of unstable zeros of $|R|$.

Two examples are given for the first flight stages of two Saturn V vehicles at 120 seconds of flight time: the AS-502 which is linearly unstable at 5 Hz and the AS-504 which is stabilized by a helium accumulator at the lox pump inlets of the outboard engines (Figures 6 and 7). The "a" figures are the nominal cases, while the "b" figures show the plots for half of the lox pump/thrust gain ($g_L/2$) which is one of the most sensitive parameters. Such variations are used to find the gain margin; e.g., by linear extrapolation we find that the 5-Hz range is approximately 5 dB unstable for AS-502, while the same frequency range becomes 6 dB stable for AS-504. The accuracy of the margin prediction can be increased by plotting new gain cases in an iterative fashion. Thus we can find margins with any degree of precision, while the plots always indicate exactly whether the system is stable or unstable.

---

*The programming effort of Mr. W. F. Crumbley of MSFC's Computation Laboratory is gratefully acknowledged.

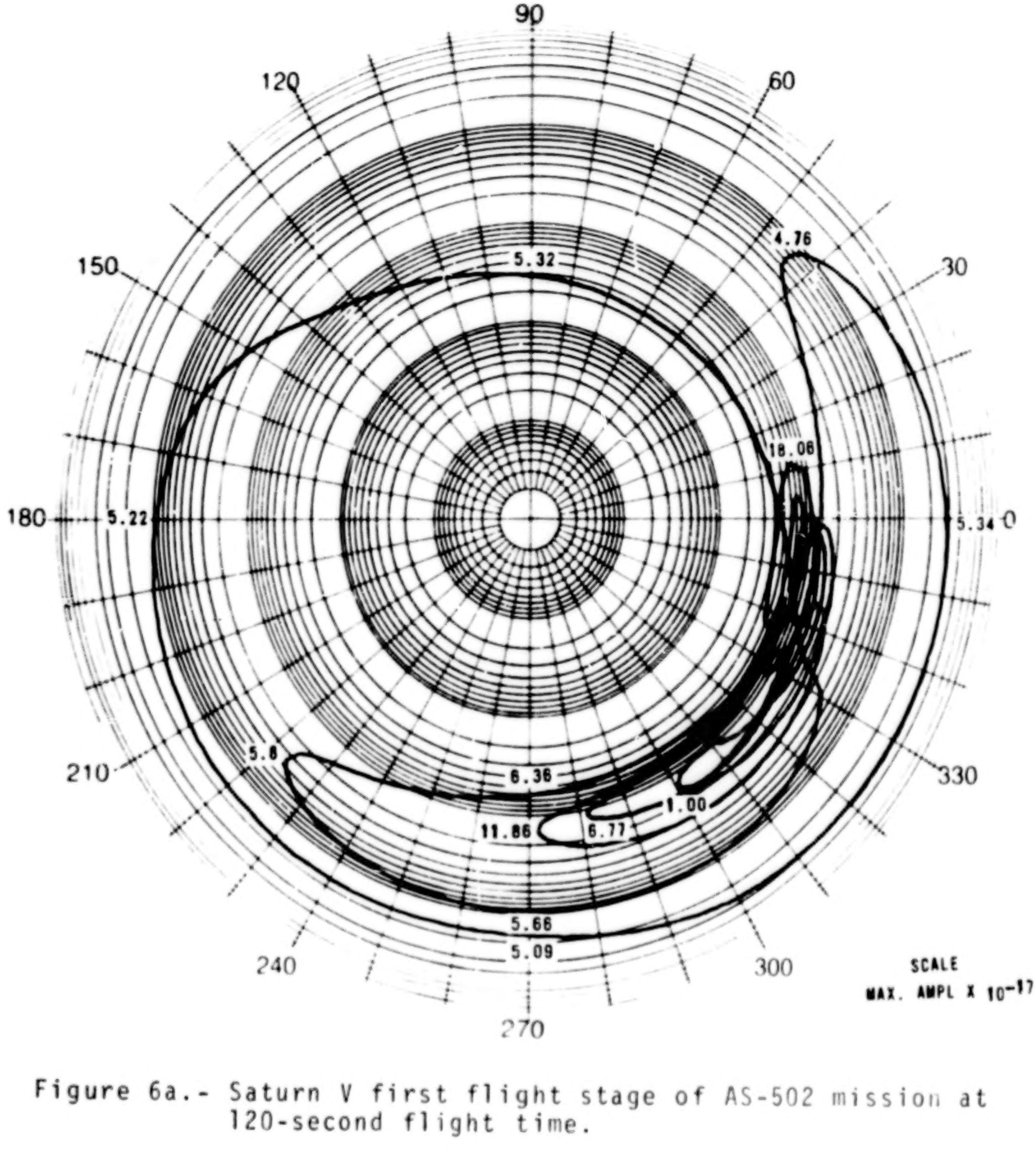


Figure 6a.- Saturn V first flight stage of AS-502 mission at 120-second flight time.

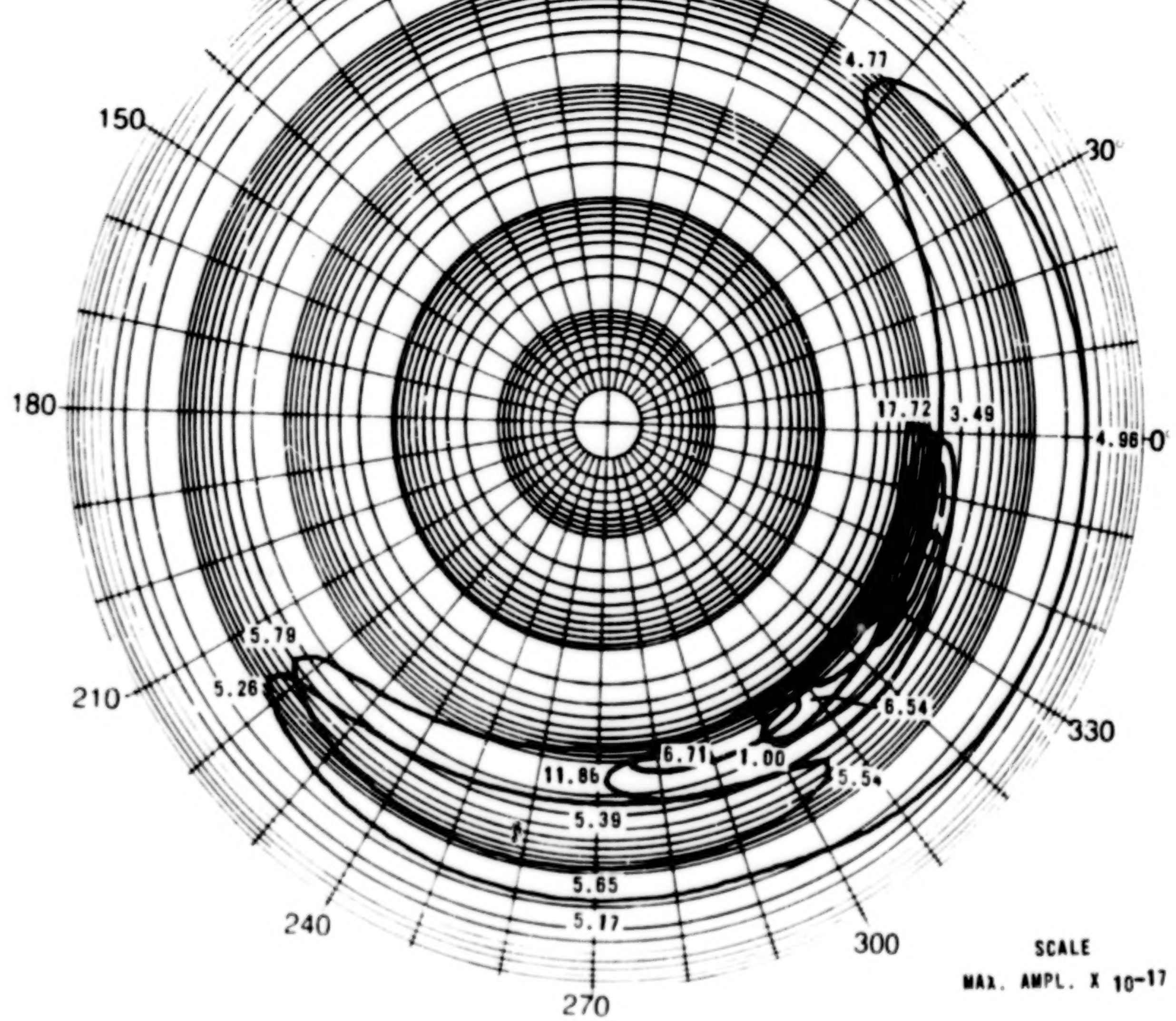


Figure 6b.- Saturn V first flight stage of AS-502 mission with one-half lox pump thrust gain.

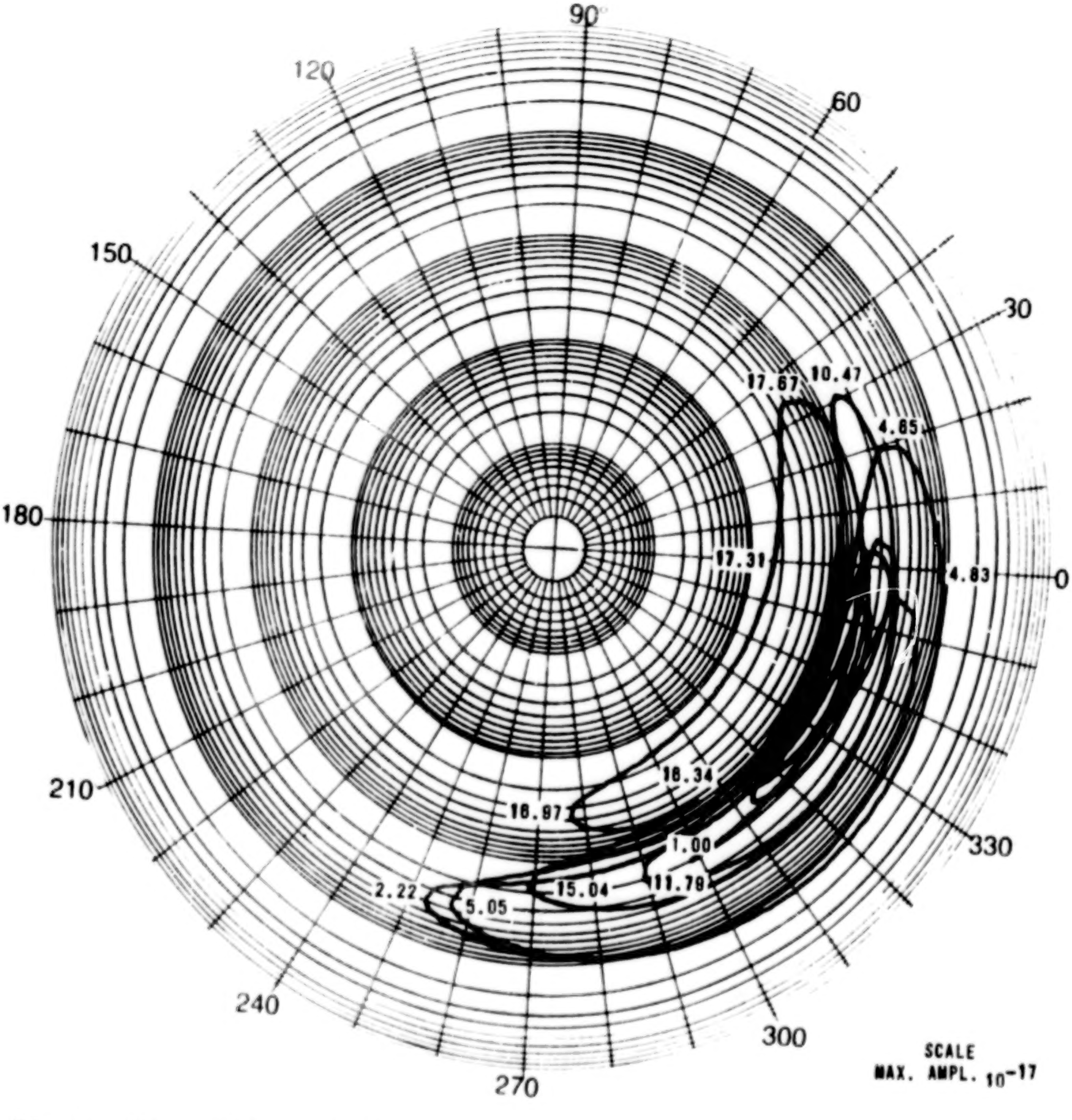


Figure 7a.- Saturn V first flight state of AS-504 mission at 120-second flight time.

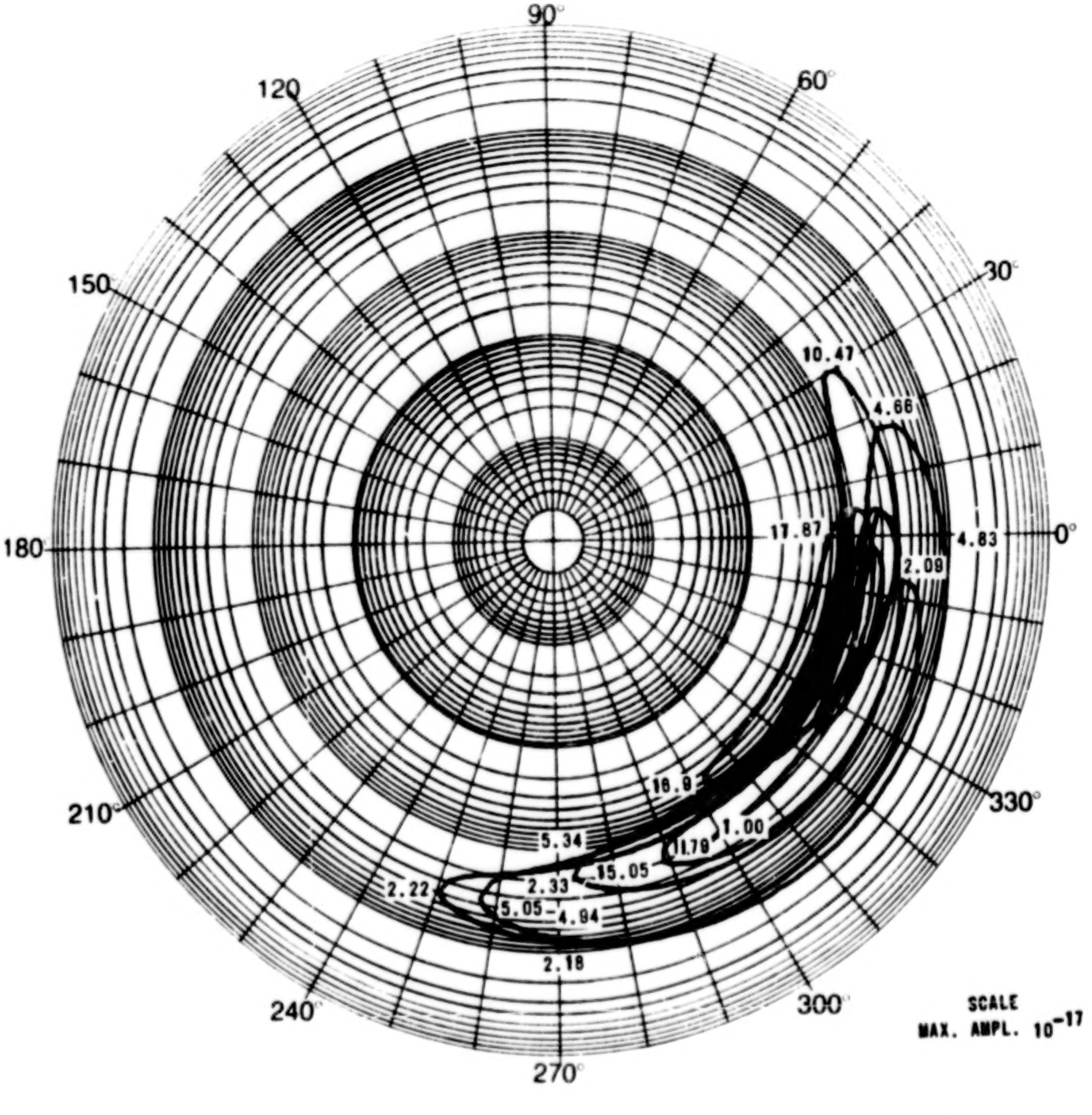


Figure 7b.- Saturn V first flight stage of AS-504 mission with one-half of lox pump/thrust gain.

One may imagine a -1 point at the plot's origin, but note that the loop gain is not a simple factor, merely one parameter or several dependent parameters which are buried in a loop-type transfer function.

CONCLUSION

All constant gain systems can be described by the stable matrix form. Stable operations always permit unlimited decompositions such that stable subsystems are obtained. The approach resembles in many respects the state-space form which is actually reached when the decomposition is carried down to the essential elements of a system. The dimensional savings are appreciable when compared with the state space which appears now as the last resort for a stable structure. The decomposition into stable subsystems rids the system of the non-controllable and non-observable parts at the subsystem's level.

The method is generally applicable to the stability analysis of vector feedback which is represented by a matrix inversion that leads to the evaluation of a determinant by a Nyquist plot about the origin. Necessary and sufficient conditions are readily derived for this method which is also well suited for the analysis of large-order systems as exemplified by the stability analysis of the pogo phenomenon of the Saturn V's first stage where eigenvalues numbered over a hundred.

REFERENCES

1. Bikle, F.: "An Investigation of Longitudinal Oscillation Instability of the Saturn V LOR Vehicle." Interim Technical Summary Report, CR-65-43. Marietta Corp., Denver, Colo., July 1965.
2. Davis, W. F., Keeton, D. L., and Lynch, T. E.: "Thor Longitudinal Oscillation Study." Report SM-45009, Douglas Missile and Space Systems Division, Santa Monica, Calif., March 1964.
3. Dent, E. J.: "Feedline Analysis for Inclusion in Pogo Stability Programs." Memorandum C1-SRL-2-001-4, Brown Engineering Company, May 1966.
4. Fashbauch, R. H., and Streeter, V. L.: "Resonance in Liquid Rocket Engine Systems." Applied Mechanics Fluid Engineering Conference (American Society of Mechanical Engineers), Washington, D. C., June 7-9, 1965.
5. Goerner, E. E.: "Lox Prevalve to Prevent Pogo Effect on Saturn V." Space/Aeronautics, December 1968.

6. Rich, R. L.: "Saturn V Pogo and A Solution." AIAA Structural Dynamics and Aeroelasticity Specialist Conference, New Orleans, La., April 16-17, 1969.

7. Rocketdyne, North American, Inc.: "Engine System Transfer Functions for Support of S-V Vehicle Longitudinal Stability (POGO) Analysis Program." Report R-6929, Canoga Park, Calif., March 28, 1967.

8. Rose, R. G., Staley, J. A., and Simson, A. K.: "A Study of System-Coupled Longitudinal Instabilities in Liquid Rockets, Part 1-Analytical Model." Report by General Dynamics/Convair, San Diego, Calif., September 1965.

9. Rubin, S.: "Longitudinal Instability of Liquid Rocket Due to Propulsion Feedback (POGO)." Journal of Spacecraft and Rockets, Volume 3, August 1966.

10. Worlund, A. L., Glasgow, V. L., Norman, D. E., and Hill, R. D.: "The Reduction of POGO Effects by Gas Injection." AIAA Second Propulsion Joint Specialist Conference, Colorado Springs, Colo., June 13-17, 1966.

11. von Pragenau, G. L.: "Stability Analysis of Apollo-Saturn V Propulsion and Structure Feedback Loop." AIAA Guidance, Control, and Flight Mechanics Conference, Princeton, N. J., August 18-20, 1969.

12. Bellman, R.: Stability Theory of Differential Equations. McGraw-Hill Book Company, Inc., New York, N. Y., 1953.

13. Bellman, R.: Introduction to Matrix Analysis. McGraw-Hill Book Company, Inc., New York, N. Y., 1960.

14. Elgerd, O. I.: Control System Theory. McGraw-Hill Book Company, Inc., New York, N. Y., 1967.

15. Gantmacher, F. R.: The Theory of Matrices. Chelsea Publishing Co., New York, N. Y., 1959.

16. Hahn, W.: "Eine Bemerkung zur zweiten Methode von Liapunov." Math. Nachr., pp. 349-354, 1956.

17. Hahn, W.: Theory and Application of Liapunov's Direct Method. Prentice Hall, Inc., 1963.

18. Hsu, J. C., and Meyer, A. U.: Modern Control Principles and Applications. McGraw-Hill Book Co., New York, N. Y., 1968.

19. Koppel, L. B.: Introduction to Control Theory with Applications to Process Control. Prentice Hall, Inc., 1968.

20. Schultz, D. G. and Melsa, T. L.: State Functions and Linear Control Systems. McGraw-Hill Book Co., Inc., New York, N. Y., 1967.

21. Timothy, L. K. and Bona, B. E.: State Space Analysis. McGraw-Hill Book Co., Inc., New York, N. Y., 1968.

22. Zadeh, L. A. and Desoer, C. A.: Linear System Theory. McGraw-Hill Book Co., Inc.

23. Kalman, R. E.: "Mathematical Description of Linear Dynamic Systems." Journal of the Society for Industrial and Applied Mathematics Series A Control, Volume 1, pp. 152-192, 1962-1963.

24. Gibson, J. E.: Nonlinear Automatic Control, McGraw-Hill Book Co., Inc., 1963.

25. Freeman, H.: "Stability and Physical Realizability Considerations in the Synthesis of Multipole Control Systems." Trans. AIEE (Applications and Industry), Volume 77, pp. 1-5, March 1958.

26. Mason, S. J.: "Feedback Theory - Further Properties of Signal Flow Graphs." Proceedings IRE, pp. 902-926, July 1956.

27. Schwarz, H.: Mehrfachregelung, Grundlagen einer Systemtheorie. Springer Verlag, New York, 1967.

28. Truxal, T. G.: Control Engineers' Handbook. McGraw-Hill Book Co., Inc., 1958.

29. Chen, C. T.: "Stability of Linear Multivariable Feedback Systems." Proceedings of the IEEE, Volume 56, No. 5, May 1968.

30. Chen, C. T.: "Representation of Linear Time Invariant Composit Systems." IEEE Transactions on Automatic Control, Volume AC-13, No. 3, pp. 227-285, June 1968.

31. Gilbert, E. G.: "Controllability and Observability in Multivariable Control Systems." J.S.I.A.M. Control Ser. A, Volume 2, No. 1, pp. 128-151, 1962-1963.

32. Newton, G. C., Jr., Gould, L. A., and Kaiser, J. F.: Analytical Design of Linear Feedback Control. John Wiley and Sons, Inc., pp. 302-309, 1961.

33. Nyquist, H.: "Regeneration Theory." Bell System Technical Journal, Volume 11, pp. 126-147, 1932, republished by Bellman, R., and Kalaba, R.: Mathematical Trends in Control Theory. Dover Publication, Inc., New York, 1946.

34. von Pragenau, G. L.: "A Hydraulic Support for Free-Flight Simulation with the Saturn V Apollo Vehicle." AIAA/JACC Guidance and Control Conference, Seattle, Wash., pp. 413-440, August 15-17, 1966.

BLANK PAGE

SYNTHESIS OF FINITE SETTLING TIME DISCRETE SYSTEMS

By James A. Gatlin
Goddard Space Flight Center

SUMMARY

This paper presents a new finite settling time (FST) control law which has the form

$$-u_{k+1} = b_0(k)u_k + \sum_{i=1}^{n} b_i(k)\theta_i(k) \qquad (1)$$

where

U_k = control level, $kT \leq t < (k+1)T$,

$\theta_i(k)$ = output error sampled at $t = (k + i/n)T$,

$b_0(k)$, $b_i(k)$ = controller coefficients, $i = 1,2,\cdots,n.$ $k = 0,1,2,\cdots$

The paper demonstrates the performance and the flexibility of this dual-rate, sampled-data law as applied to the class of error driven, feedback control systems illustrated in Figure 1. The performance is superior to that obtained by using the FST control law for discrete systems discussed by Kalman and Bertram (ref. 1) and Lindorff (ref. 2). Their procedure is based upon the availability of all the state variables and a plant with a rational transfer function, and the design obtained suffers from high noise sensitivity and restricted dynamic range due to plant saturation.

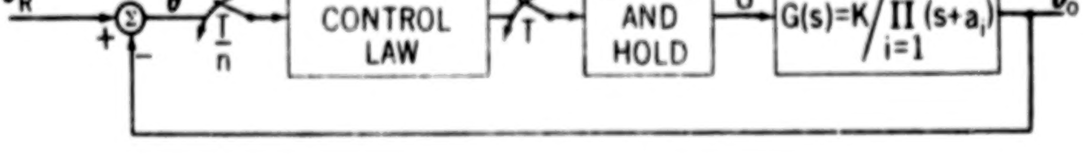


Figure 1.- A dual rate sampled data system

The new control law serially accumulates n-weighted, plant output error samples – $\theta_i(k)$, $i = 1, . . ., n$ – during each T-second interval. Each sample is separated by T/n seconds with the first sample taken at $t = kT + (T/n)$. At the end of the k-th interval – that is, at $t = (k + 1)T$ – the construction of a new control level is completed. Then retaining only the value of the previous control level, $b_0(k)U_k$, the control law construction cycle repeats.

The synthesis method involves selection of the coefficients $b_i(k)$, $i = 1, . . ., n$ such that the system is brought to a stable null equilibrium in m + 1 control level updates where m is the order of the plant. Also when $b_i(k) = b_i$ – that is, the value of b_i is constant for k control law construction cycles – the synthesis method minimizes

$$F = \sum_{i=1}^{n} \left(b_i\right)^2$$

in order to minimize input noise transmission.

A hybrid analog/digital simulation is performed in order to illustrate the practical aspects of this type of finite settling time design.

FORMULATION OF THE PROBLEM

The differential equation describing the plant response can be written as

$$\dot{x}(t) = Gx(t) + gu(t) \tag{2}$$

$$x(0) = x_0$$

where G is a m×m matrix, x(t) is the plant m-state vector, x_0 defines the plant initial state, and g is the driving vector which for this single input u(t) has all elements zero except the last which is the control gain constant K. Equation (2) has the solution

$$x(t) = \Phi(t)\, x_0 + \Phi(t) \int_0^t \Phi(-\tau)\, gu(\tau)\, d\tau \tag{3}$$

where $\Phi(t) = e^{Gt}$ is the plant transition matrix. Letting

$$x_k(\ell) = x(t)\big|_{t=k+\ell T/n}, \qquad \ell = 1, 2, . . ., n \tag{4}$$

and

$$u(t) = u_k, \qquad kT \le t < (k+1)T$$

Eq. (3) yields the state vector at the error sample times during the k-th interval as

$$x_k(\ell) = \Phi\left(\frac{\ell T}{n}\right) x_k(0) + \Phi\left(\frac{\ell T}{n}\right) \int_0^{\ell T/n} \Phi(-\tau)\ g u(\tau)\ d\tau \qquad (5)$$

Selecting the first element of x(t) to be the plant output variable θ yields the output samples

$$\theta_k(\ell) = \phi_1^T \left(\frac{\ell T}{n}\right) x_k(0) + \phi_1^T \left(\frac{\ell T}{n}\right) \left[h\left(\frac{\ell T}{n}\right)\right] u_k \qquad (6)$$

where ϕ_1^T is the first row of the matrix Φ and

$$h\left(\frac{\ell T}{n}\right) = \int_0^{\ell T/n} \Phi(-\tau)\ g\ d\tau \qquad (7)$$

In what follows $\Phi(\ell T/n)$ will be written as $\Phi(\ell)$ and $\phi_1^T(\ell T/n)h(\ell T/n)$, a scalar depending only on ℓ, will be represented by $\alpha(\ell)$. For $\ell = 1,2,\ldots,n$ Eq. (6) produces

$$\begin{bmatrix} \theta_k(1) \\ \vdots \\ \theta_k(n) \end{bmatrix} = \begin{bmatrix} \phi_1^T(1) \\ \vdots \\ \phi_1^T(n) \end{bmatrix} x_k(0) + \begin{bmatrix} \alpha(1) \\ \vdots \\ \alpha(n) \end{bmatrix} u_k \qquad (8)$$

Defining

$$y_k^T = \left[\theta_k(1) \ldots \theta_k(n)\right], \qquad \text{n-vector}$$

$$V^T = \left[\phi_1^T(1) \ldots \phi_1^T(n)\right], \qquad n \times m \text{ matrix}$$

$$\alpha^T = \left[\alpha(1) \ldots \alpha(n)\right], \qquad \text{n-vector} \qquad (9)$$

reduces Eq. (8) to

$$y_k = Vx_k(0) + \alpha\, u_k \tag{10}$$

Likewise Eq. (1) can be expressed in vector notation as

$$-u_{k+1} = b_0 u_k + b^T y_k \tag{11}$$

where $b^T = \left[b_1(k) \ldots b_n(k) \right]$. Combining Eqs. (10) and (11) yields

$$-u_{k+1} = b^T V x_k(0) + \left(b_0 + b^T \alpha \right) u_k \tag{12}$$

In Eq. (12) the output samples, $\theta_i(k)$, have been replaced by a function of the state of the plant at the beginning of the k-th control interval, $x_k(0)$, and control level, u_k, applied during the interval. Let c, an m vector, be defined as

$$c^T = b^T V \tag{13}$$

or

$$c = V^T b$$

so that Eq. (12) also has the form

$$-u_{k+1} = c^T\, x_k(0) + c_0 u_k \tag{14}$$

where $c_0 = b_0 + b^T \alpha$.

The control vector c_0, c^T can be synthesized from the weighting vector b_0, b^T as follows. For $n = m$

$$b = \left[V^T \right]^{-1} c$$

$$b_0 = b_0 - b^T \alpha \tag{15}$$

and for $n \geq m + 1$, b_0 can be left as a free parameter so that

$$C_{m+1} = \begin{bmatrix} c \\ - - - - \\ c_0 \quad - b_0 \end{bmatrix} = \begin{bmatrix} V^T \\ -- \\ \alpha \end{bmatrix} \quad b = Hb \tag{16}$$

Thus for $n = m+1$

$$b\left(b_0\right) = H^{-1} C_{m+1} \tag{17}$$

and for $n>m+1$, the solution of Eq. (16) that also minimizes

$$F = \sum_{i=1}^{n} {b_i}^2 \tag{18}$$

is

$$b_{OPT}\left(b_0\right) = H^T \left[HH^T\right]^{-1} C_{m+1} \tag{19}$$

For either Eq. (17) or (19) c_0 may be arbitrarily selected or selected so that F is also minimized with respect to c_0.

Minimization of the noise content factor, F, results in the minimum noise transmission through the controller since for white noise input,

$$\sigma_u^2 / \sigma_\theta^2 = F \tag{20}$$

where σ_u^2 is the variance of the control level sequence and σ_θ^2 is the variance of the output sample sequence.

In this section the equivalence between Eqs. (1) and (14) has been developed. In the following section the unique values of c_0, c^T required for FST response will be determined. The values of b_0, b^T used in Eq. (1) are not unique but depend on the number and the format of the error samples. Equations (15), (17), and (19) can be used to determine the values of b_0, b^T required to construct any given $(m + 1)$ control vector c_0, c^T.

DETERMINATION OF c_0 AND c^T FOR FST RESPONSE

The change in the state of the plant during the k-th interval is given by Eq. (5) with $\ell = n$:

$$x_k(n) = x_{k+1}(0) = \Phi(n)x_k(0) + \Phi(n)h(n)u_k \tag{21}$$

Equations (14) and (21) can be combined to yield

$$\begin{bmatrix} x_{k+1} \\ --- \\ u_{k+1} \end{bmatrix} = A \begin{bmatrix} x_k(0) \\ --- \\ u_k \end{bmatrix} \tag{22}$$

where

$$A = \left[\begin{array}{c|c} \Phi(n) & \Phi(n)h(n) \\ \hline -\,c^T & -\,c_0 \end{array}\right] = \left[a_{ij}\right] \tag{23}$$

The matrix A is the controller plus plant, closed-loop, state transition matrix. To obtain FST response in $(m + 1)$ steps for any set of initial conditions, the parameters c_0, c^T must be chosen so that

$$A^{m+1} \equiv 0. \tag{24}$$

Direct solution of Eq. (24) is quite tedious; but from the Caley-Hamilton theorem, Eq. (24) is satisfied if A, an $(m + 1) \times (m + 1)$ matrix, has all its eigenvalues zero. The cannonical form of such a matrix has zeros everywhere except the diagonal above the main diagonal which contains unity elements. Equation (24) is satisfied if

$$|zI-A| = z^{m+1} + \sum_{i=0}^{m} a_i z^i = 0 \tag{25}$$

has

$$a_i = 0, \qquad i = 0, 1, \cdots, m \tag{26}$$

There are $(m + 1)$ equations involved in Eq. (25) and $(m + 1)$ controller parameters c_0, c^T. In general, expansion of Eq. (25) about the last row of $|zI-A|$ followed by the conditions of Eq. (26) yields

$$Qc = q,$$

$$c_0 = \text{Trace } \Phi(A) \tag{27}$$

The matrix Q and the vector q are systematically obtained from A as illustrated below for a third order matrix:

$$\begin{vmatrix} z-a_{11} & -a_{21} & -\alpha_1 \\ -a_{12} & z-a_{22} & -\alpha_2 \\ c_1 & c_2 & z + c_0 \end{vmatrix} = 0 \tag{28}$$

Expanding about the last row yields

$$\begin{aligned} 0 = \left(\alpha_2 a_{12} - \alpha_1 a_{22} + \alpha_1 z\right) c_1 \\ + \left(\alpha_1 a_{21} - \alpha_2 a_{11} + \alpha_2 z\right) c_2 \\ + \left(a_{11} a_{22} - a_{12} a_{21} - \left(a_{11} + a_{22}\right) z + z^2\right)\left(z + c_0\right) \end{aligned} \tag{29}$$

Setting $a_2 = 0$ produces $c_0 = a_{11} + a_{22}$, and setting $a_1 = 0$, $a_0 = 0$ yields

$$Q = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \left(\alpha_2 a_{12} - \alpha_1 a_{22}\right) & \left(\alpha_1 a_{21} - \alpha_2 a_{11}\right) \end{bmatrix}$$

and

$$q = \begin{bmatrix} c_0^2 - \left(a_{11} a_{22} - a_{12} a_{21}\right) \\ - \left(a_{11} a_{22} - a_{12} a_{21}\right) c_o \end{bmatrix} \tag{30}$$

DESIGN EXAMPLE

Consider the system of Figure 1 with $m = 2$, $a_1 = 0$, and $a_2 = a$. Given the parameters K, a, T, and $n \geq 2$, the controller parameters $b_o b^T$ are to be chosen to yield FST transient response and a minimum noise content factor, F.

The fictitious controller parameters c_o, c^T, illustrated in Figure 2, are obtained first. Letting $b = e^{-aT}$ results in

$$\Phi(n) = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} = \begin{bmatrix} 1 & 1-b/a \\ 0 & b \end{bmatrix} \tag{31}$$

and

$$h(n) = \frac{K}{a} \begin{bmatrix} \left(1 + aT - b^{-1}\right)/a \\ \left(b^{-1} - 1\right) \end{bmatrix}$$

with

$$\Phi(n)h(n) = \frac{K}{a} \begin{bmatrix} (aT + b - 1)/a \\ (1 - b) \end{bmatrix} \tag{32}$$

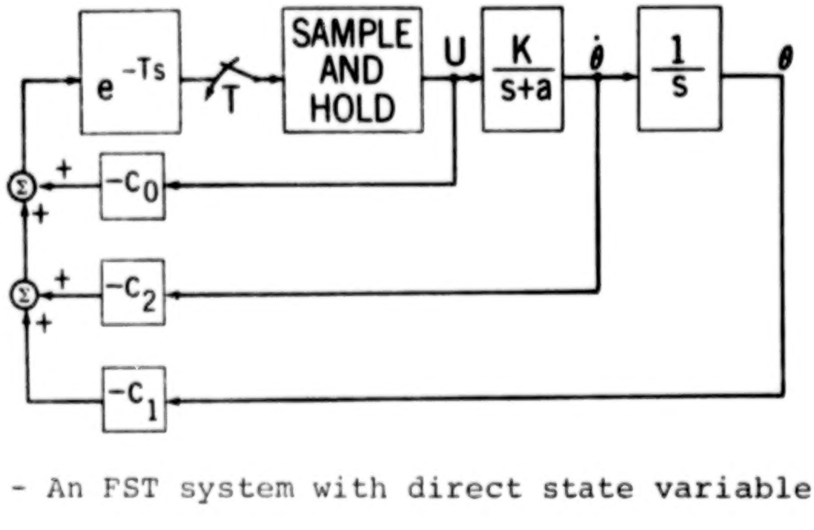


Figure 2. - An FST system with direct state variable feedback

As defined in Eq. (23), the matrix A is

$$A = \left[\begin{array}{cc|c} 1 & \frac{(1 - b)}{a} & \frac{(aT + b-1)}{(aT)^2} \; KT^2 \\ 0 & b & \frac{(1 - b)}{aT} \; KT \\ \hline -c_1 & -c_2 & -c_0 \end{array}\right] \tag{33}$$

and using Eq. (30) yields

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = Q^{-1}q = \begin{bmatrix} \dfrac{aT}{(1 - b)KT^2} \\ \dfrac{1 - b - aTb^3}{(1-b)^2 \; KT} \end{bmatrix} \tag{34}$$

with $c_0 = 1 + b$. It is sometimes useful to set $c_1 = 1$; that is, to use unity static position feedback. This allows the FST loop gain to be expressed as

$$K_0T^2 = \frac{aT}{(1 - b)} \tag{35}$$

and reduces the expression for c_2 to

$$c_2 = \frac{1 - b - aTb^3}{(1 - b) \; a} \tag{36}$$

For n = 2 Eq. (15) yields the weighting parameters b_0, b_1, b_2 from the calculated fictitious controller parameters c_0, c_1, c_2. Using

$$V^T = \begin{bmatrix} 1 & 1 \\ \dfrac{1 - b^{0.5}}{a} & \dfrac{1 - b}{a} \end{bmatrix}$$

$$\alpha = \frac{kT^2}{(aT)^2} \begin{bmatrix} aT/2 + b^{0.5} - 1 \\ aT + b - 1 \end{bmatrix} \tag{37}$$

consider the case for $aT \sim 0$. Since $e^{-x} \doteq 1-x+x^2/2!-x^3/3!+\cdots$, the general expressions reduce to

$$K_oT^2 = aT\Big/\left(aT - \frac{1}{2}(aT)^2 + \cdots\right) \to 1$$

$$c_2 = \frac{1}{a}\left(\frac{aT - \frac{1}{2}(aT)^2 + \cdots - aT + 3(aT)^2 - \cdots}{aT - \frac{1}{2}(aT)^2 + \cdots}\right) \to \frac{5T}{2} \qquad (38)$$

with $c_o = 2$. Equation (15) now is

$$b = \begin{bmatrix} 1 & 1 \\ T/2 & T \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 5T/2 \end{bmatrix} = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$$

and

$$b_0 = 2 - [-3 \quad 4] \begin{bmatrix} 1/8 \\ 1/2 \end{bmatrix} = 3/8 \qquad (39)$$

For $n = 2$ the noise content factor is

$$F = (4)^2 + (3)^2 = 25 \qquad (40)$$

Now consider the use of $n = 3$. The fictitious controller parameters, for $KT^2 = 1$, are

$$C_{m+1} = \begin{bmatrix} c_1 \\ c_2 \\ c_0 - b_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 5T/2 \\ 2 - b_0 \end{bmatrix} \qquad (41)$$

and as defined in Eq. (16),

$$H = \begin{bmatrix} 1 & 1 & 1 \\ T/3 & 2T/3 & 3T/3 \\ 1/18 & 4/18 & 9/18 \end{bmatrix} \qquad (42)$$

Equation (17) yields

$$b_1 = \frac{9}{4} - 9b_0$$

$$b_2 = -9 + 18b_0$$

$$b_3 = \frac{31}{4} - 9b_0 \tag{43}$$

Calculating the noise content factor and $dF(b_0)/db_0 = 0$ produces

$$(b_0)_{OPT} = -\left(\begin{bmatrix} \frac{9}{4} & -9 & \frac{31}{4} \end{bmatrix}\begin{bmatrix} -9 \\ 18 \\ -9 \end{bmatrix}\right) \Bigg/ \left(\begin{bmatrix} -9 & 18 & -9 \end{bmatrix}\begin{bmatrix} -9 \\ 18 \\ -9 \end{bmatrix}\right) \tag{44}$$

resulting in $(b_0)_{OPT} = 0.5185$, $F_{MIN} = 15.5$, and

$$b_{OPT} = \begin{bmatrix} -2.417 \\ 0.333 \\ 3.083 \end{bmatrix} \tag{45}$$

A unique feature of the general control law expressed by Eq. (1) is the incorporation of memory of the previous control level. If $b_0 = 0$ is used, FST response is obtained at the expense of $F = 146$. The low noise transmission feature of the disclosed FST control law is one of its most attractive features.

For $n > m + 1$ the matrix H is $(m + 1) \times n$. Equation (19) yields, for $n = 4$,

$$b_{OPT} = \frac{1}{4}\begin{bmatrix} 11 - 32\,b_0 \\ -21 + 32\,b_0 \\ -15 + 32\,b_0 \\ 29 - 32\,b_0 \end{bmatrix} = \begin{bmatrix} -2 \\ -0.5 \\ 1 \\ 2.5 \end{bmatrix} \tag{46}$$

with $(b_0)_{OPT} = 0.5938$ and $F_{MIN} = 11.5$.

For $n > m + 1$ it is sometimes impractical to use n distinct coefficients. Using only four, Eq. (1) takes the form

$$-u_{k+1} = b_0 u_k + b_1 \sum_{i=1}^{n/4} \theta_i(k) + b_2 \sum_{i=n/4+1}^{n/2} \theta_i(k) + b_3 \sum_{i=n/2+1}^{3n/4} \theta_i(k) + b_4 \sum_{i=3n/4+1}^{n} \theta_i(k) \qquad (47)$$

With the error samples thus grouped in 4 sets of n/4 samples, the matrix H is

$$H = \begin{bmatrix} \frac{n}{4} & \frac{n}{4} & \frac{n}{4} & \frac{n}{4} \\ \frac{n+4}{32} T & \frac{3n+4}{32} T & \frac{5n+4}{32} T & \frac{7n+4}{32} T \\ \frac{n^2+6n+8}{384n} & \frac{7n^2+18n+8}{384n} & \frac{19n^2+30n+8}{384n} & \frac{37n^2+42n+8}{384n} \end{bmatrix} \qquad (48)$$

and for large n, H becomes

$$H = \left(\frac{n}{4}\right) \begin{bmatrix} 1 & 1 & 1 & 1 \\ T/8 & 3T/8 & 5T/8 & 7T/8 \\ 1/96 & 7/96 & 19/96 & 37/96 \end{bmatrix} \qquad (49)$$

resulting in

$$\left(\frac{n}{4}\right) b_{OPT} = \begin{bmatrix} -2.150 \\ -0.550 \\ 1.050 \\ 2.650 \end{bmatrix} \qquad (50)$$

with $(b_o)_{OPT} = 0.8333$ and $F_{MIN} = (13.05)\ (4/n)$.

As $n \to \infty$ Eqs. (47) and (49) become

$$-u_{k+1} = b_o u_k + b_1 \int_0^{T/4} \theta(k,\ t)\ dt + b_2 \int_{T/4}^{T/2} \theta(k,\ t)\ dt$$

$$+ b_3 \int_{T/2}^{3T/4} \theta(k,\ t)\ dt + b_4 \int_{3T/4}^{T} \theta(k,\ t)\ dt \tag{51}$$

and

$$H = \left(\frac{T}{4}\right) \begin{bmatrix} 1 & 1 & 1 & 1 \\ T/8 & 3T/8 & 5T/8 & 7T/8 \\ 1/96 & 7/96 & 19/96 & 37/96 \end{bmatrix} \tag{52}$$

with $(T/4)\ b_{OPT} = (n/4)\ b_{OPT}$. The calculation of F_{MIN}, however, must now include the description of the noise source power spectrum. As an example, consider a band-limited noise generator with uniform power density,

$$P_d = \sigma_0^2 \big/ f_0, \qquad -f_0 < f < f_0 \tag{53}$$

passed through a linear, low-pass filter with the Laplace transfer function

$$G(s) = \frac{\sqrt{\lambda}}{s + \lambda}, \qquad \lambda << 2\pi f_0 \tag{54}$$

The resulting noise source is described by the variance

$$\sigma_e^2 = \frac{1}{j^{2\pi}} \int_{-j\infty}^{j\infty} \frac{P_d}{2} G(s)G(-s)\ ds = \frac{1}{4} \left(\sigma_0^2 \big/ f_0\right) \tag{55}$$

and the auto-correlation function,

$$R_{NN}(\tau) = \sigma_e^2\ e^{-\lambda\tau} \tag{56}$$

For samples taken every $T/n<1/\lambda$ seconds, it is no longer appropriate to assume statistical independence of the samples; however, if the noise source is integrated for $T/4>1/\lambda$ seconds, samples taken every T/4 seconds can be assumed independent with

$$\sigma_I^2 = \int_0^{T/4} \left(\frac{T}{4} - \tau\right) R_{NN}(\tau)\, d\tau \tag{57}$$

which yields

$$\frac{\sigma_I^2}{\sigma_e^2} = \frac{T}{4\lambda} \tag{58}$$

and

$$F = \frac{\sigma_u^2}{\sigma_e^2} = \frac{\sigma_u^2}{\sigma_I^2} \times \frac{\sigma_I^2}{\sigma_e^2} = \left[b_4^2 + b_3^2 + b_2^2 + b_1^2 \right] \frac{T}{4\lambda}$$

or

$$F = \left[\left(\frac{T}{4} b_4\right)^2 + \left(\frac{T}{4} b_3\right)^2 + \left(\frac{T}{4} b_2\right)^2 + \left(\frac{T}{4} b_1\right)^2 \right] \Big/ \left(\frac{4\lambda}{T}\right)\left(\frac{T}{4}\right)^2 \tag{59}$$

Therefore as for Eq. (50), $F_{MIN} = (13.05)\ (4/\lambda T)$.

The preceding cases show how F is reduced for $n>m$; in summary Table I* is given.

TABLE I

n	$(b_o)_{OPT}$	$\sqrt{F_{MIN}}$	$\lambda T/n$
2	0.3750	5.000	1024
3	0.5185	3.937	683
4	0.5938	3.391	512
32	0.8022	1.267	64
128	0.8255	0.637	16
512	0.8314	0.319	4
∞	0.8333	0.160	0

---

*for noise with $R_{NN} = e^{-\lambda\tau}, \lambda T = 2048$.

SYSTEM PERFORMANCE

Construction of an FST system to control a second order plant was accomplished by using hybrid simulation. The continuous plant, $G(s) = K/S^2$, was modeled on the analog computer, and digital computation generated the control law of Eq. (1) and the output error

$$\theta = \theta_0 - \theta_R \tag{60}$$

where θ_0 is the plant output and θ_R is a staircase reference function that changes every T seconds.

The control law update interval, T, is 11 seconds of which 10 seconds is used for output sampling and 1 second for the computational and conversion delays; that is Eq. (1) has the form

$$-u_{k+1} = b_0 u_k + b_1 \sum_{i=1}^{n/4} \theta_i(k) + b_2 \sum_{i=n/4+1}^{n/2} \theta_i(k) + b_3 \sum_{i=n/2+1}^{3n/4} \theta_i(k)$$

$$+ b_4 \sum_{i=3n/4+1}^{n} \theta_i(k) + b_5 \sum_{i=n}^{11/\Delta T} \theta_i(k) \tag{61}$$

with $b_5 = 0$ and

$$n = \left(T - T_d\right) / \Delta T = \frac{10}{\Delta T} \tag{62}$$

Output error sampling periods, ΔT, of 10/512, 10/128, and 10/32 seconds are provided; and an option that allows the plant output to be integrated for 10/4 seconds on the analog computer and sampled every 10/4 seconds is included.

Since the data sampling interval is now only $(T - T_d)$, Eq. (52) becomes

$$\begin{bmatrix} c_1 \\ c_2 \\ c_0 - b_0 \end{bmatrix} = \left(\frac{n}{4}\right) \begin{bmatrix} 1 & 1 & 1 & 1 \\ \frac{(T-T_d)}{8} & \frac{3(T-T_d)}{8} & \frac{5(T-T_d)}{8} & \frac{7(T-T_d)}{8} \\ \frac{(T-T_d)^2/T^2}{96} & \frac{7(T-T_d)^2/T^2}{96} & \frac{19(T-T_d)^2/T^2}{96} & \frac{37(T-T_d)^2/T^2}{96} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \tag{63}$$

and letting $d = T_d/(T-T_d)$, Eq. (63) reduces to

$$\begin{bmatrix} c_1 \\ (1+d)\ c_2/T \\ (1+d)^2\ (c_0-b_0) \end{bmatrix} = \left(\frac{n}{4}\right) \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1/8 & 3/8 & 5/8 & 7/8 \\ 1/96 & 7/96 & 19/96 & 37/96 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \tag{64}$$

As shown in reference 3, the relative stability as described by the root locus plot (see Figure 3) depends only on the plant gain, K, and the value of b_o. If b_o is kept at the value of $(b_o)_{opt}$ for d = 0, the addition of the delay affects only the value of F. The inverse of Eq. (64) yields

$$\left(\frac{n}{4}\right) b = \begin{bmatrix} -2.150 \\ -0.550 \\ 1.050 \\ 2.650 \end{bmatrix} + d \begin{bmatrix} 5.667 \\ -9.667 \\ -7.667 \\ 11.667 \end{bmatrix} + \left(\frac{28}{3}\right) d^2 \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \tag{65}$$

for $b_0 = 0.8333$. For d = 1/10

$$\left(\frac{n}{4}\right) b = \begin{bmatrix} -1.490 \\ -1.610 \\ 0.190 \\ 3.910 \end{bmatrix} \tag{66}$$

with F = (20.13) (4/n) which corresponds to an increase in F of 54 percent.

As developed previously - see Eqs. (35) and (38) - the FST loop gain requires a plant constant, K_0, such that

$$K_0 = \frac{1}{T^2} = \frac{1}{(11)^2} = \frac{1}{121}\ \frac{\text{volts}}{\text{sec}^2} \tag{67}$$

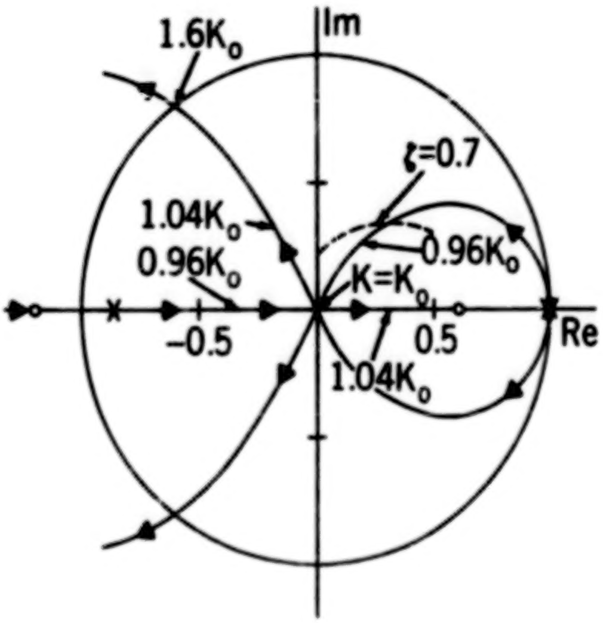
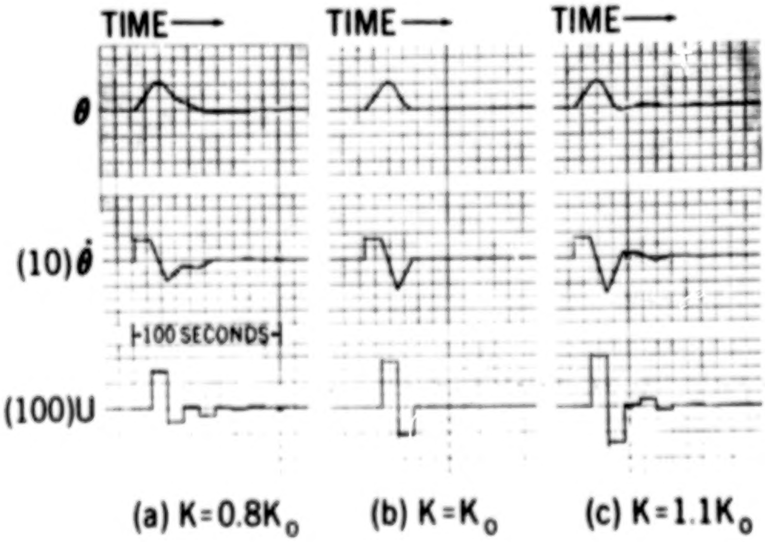


Figure 3.- FST root locus for $b_0 = 0.8333$

If $K \neq K_o$, a non-finite but rapidly convergent sequence will characterize the transient response. Figure 4 presents this effect for an initial condition response, $\dot{\theta}_o \neq 0$, $\theta_o = 0$; and Figures 5, 6, 7 show this effect when the controller is tracking an input "staircase" sinusoid.



(a) $K = 0.8K_0$ (b) $K = K_0$ (c) $K = 1.1K_0$

Figure 4.- Initial rate transient response

The coefficient set of Eq. (66) is exact for $n \to \infty$ and is a good approximate set for large "n". Figure 8 shows the effect of the approximation for n = 32, n = 128, and n = 512; and Figure 9 compares the sampling format for finite and infinite "n".

The steady-state performance when output noise is present shows, in Figure 10, the reduction of the controller noise sensitivity as the error sampling rate is increased. The runs in Figure 10 correspond to the values of n = 32, 128, 512, and ∞ given in Table I of the preceding section. Figure 11 presents a transient response with noise present for n = 512.

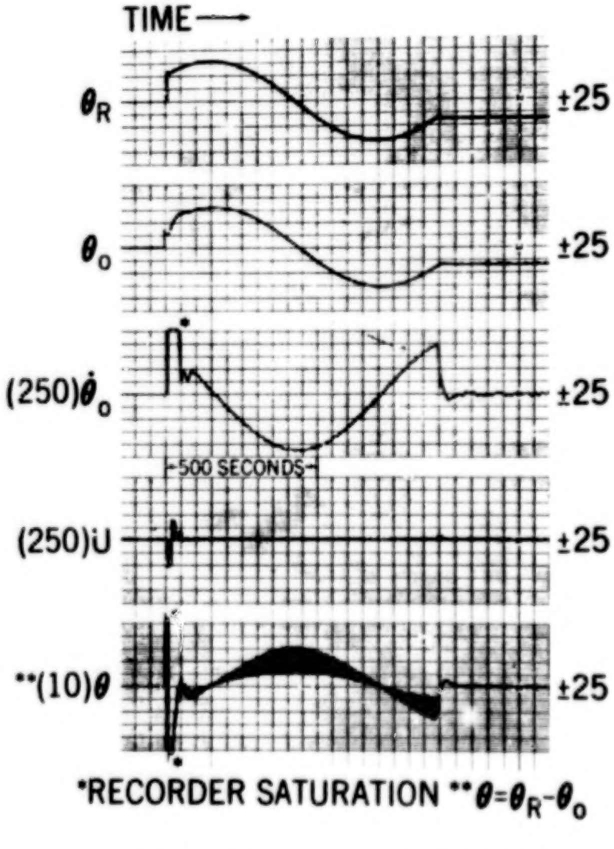


Figure 5.- Tracking response for $K = 0.8\ K_0$

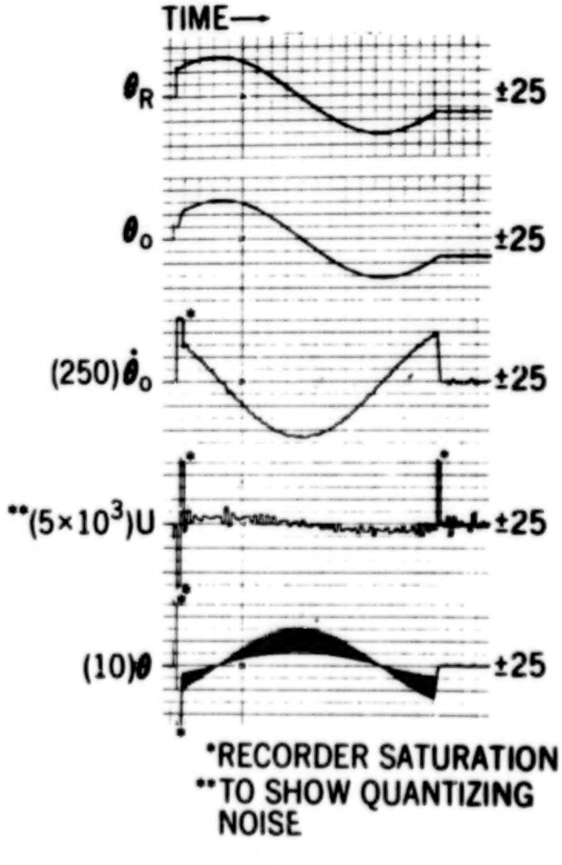


Figure 6.- Tracking response for $K = K_0$

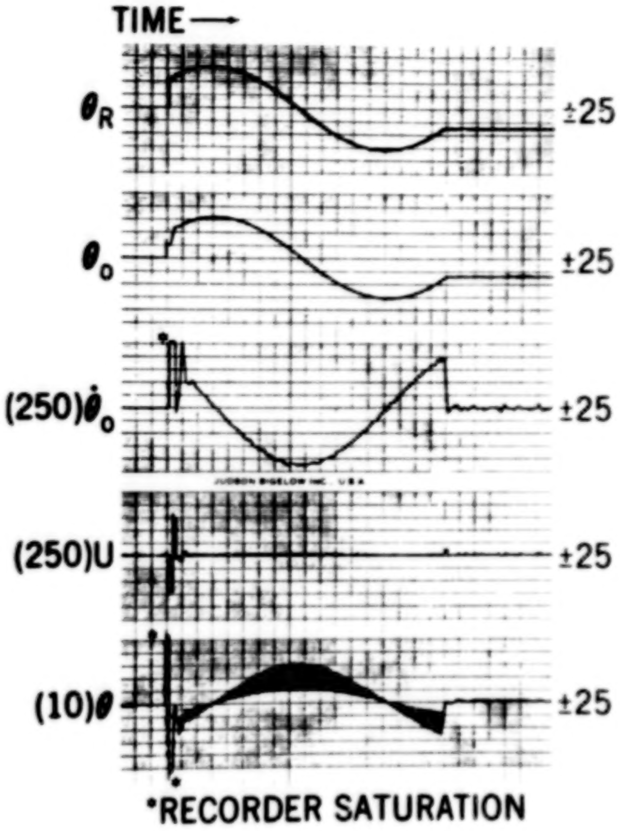
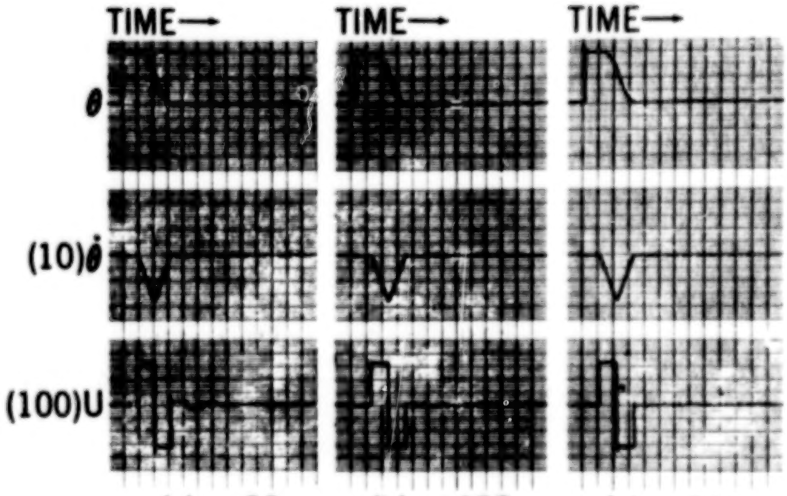
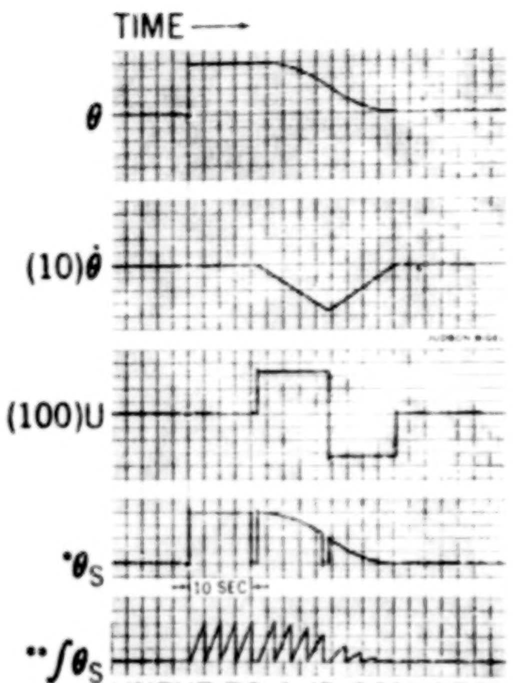


Figure 7.- Tracking response for $K = 1.1\ K_0$



(a) n=32 (b) n=128 (c) n=512

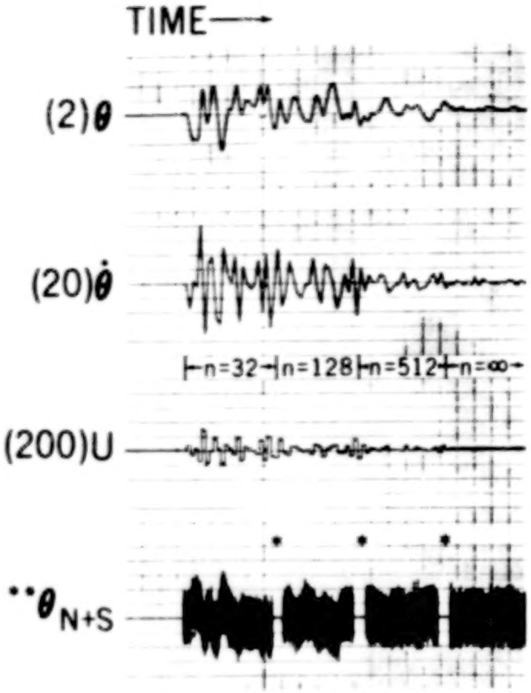
Figure 8.- The effect of the large n assumption



*INPUT TO A/D CONVERTER FOR FINITE n

**INPUT TO A/D CONVERTER FOR INFINITE n

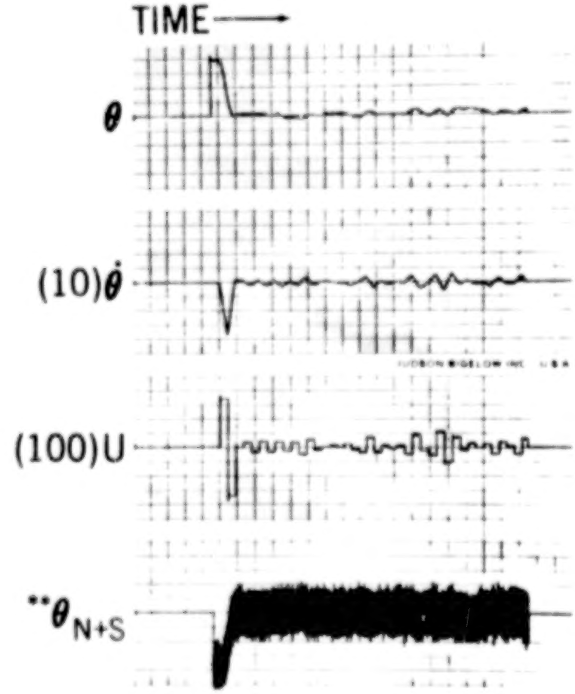
Figure 9.- Error sampling format



*RECORDER CHANNEL GAIN=ZERO

$^{**}\theta_{N+S}$=(noise-θ)

Figure 10.- Steady state performance with noise



$^{**}\theta_{N+S}$=(noise-θ)

Figure 11.- Transient performance with noise

CONCLUSIONS

This paper has extended the work done in reference 3 so that the dual-rate FST control law can be applied to n-th order single-input, single-output systems. Also one means of handling computation and conversion delays, additional noise analysis, and additional hybrid simulation work is documented.

The general control law formulation given in Eq. (1) allows for the controller coefficients $b_o(k)$, $b_i(k)$ to be a function of the control interval index, k. In the design example presented in this paper, the coefficients required are constant. The need for varying coefficients arises when the matrix A (see Eq. (23)) has elements that change from control interval to control interval. This requires that the control vector c_o, c^T also be a function of "k" if A is to approach an FST matrix as its elements approach a constant value. This approach has been used to apply the FST control law to pulse-width-modulation control of a second order plant. The elements of A can also change because of changes in the equilibrium position of a non-linear plant. Thus using the non-linear model, the coefficients $b_o(k)$, $b_i(k)$ can be adjusted to "track the operating point" and maintain FST response about the operating point.

In summary the key features of the new FST control law are:

1. The dual error sampling/control level update rates allow the error sampling rate to be chosen to meet noise sensitivity requirements and the control update rate chosen to avoid plant saturation.
2. The memory of the previous control level greatly reduces the need for the controller coefficients to produce derived error rate information.
3. Only the output state need be directly sensed; all the state variables appear combined in the generated control level, but their isolation is not required.
4. The control technique possesses null stability and reasonable tolerance to errors in the plant model.
5. The finite memory of the control law allows the controller weighting coefficients to be readily changed during each control interval.
6. A multiplicity of error samples taken while the control level remains constant provides a source of data for plant and disturbance on-line modeling.

ACKNOWLEDGMENT

The author wishes to acknowledge and thank Dr. N. N. Puri of Rutgers University, New Brunswick, New Jersey for his interest and assistance in preparing parts of this paper.

REFERENCES

1. Kalman, R. E., and Betram, J. E.: "General Procedure for Computer Control of Single-loop and Multiloop Linear Systems." Trans. AIEE, 78(2) 1959, pp 602-609.
2. Lindorff, D. P.: Theory of Sampled-data Control Systems. John Wiley and Sons, Inc., New York, 1965.
3. Gatlin, J. A.: Development of a Simple Digital Control Law for Spin Phase Synchronization of a Cartwheel Satellite in a Circular Polar Orbit. Ph.D. Thesis, University of Maryland, 1968.

DESIGN AND PERFORMANCE OF HEART ASSIST OR ARTIFICIAL HEART CONTROL SYSTEMS

By John A. Webb, Jr., and Vernon D. Gebben
Lewis Research Center

SUMMARY

The factors leading to the design of a controlled driving system for either a heart assist pump or artificial heart are discussed. The system provides square pressure waveform to drive a pneumatic-type blood pump.

For assist usage the system uses an R-wave detector circuit that can detect the R-wave of the electrocardiogram in the presence of electrical disturbances. This circuit provides a signal useful for synchronizing an assist pump with the natural heart. It synchronizes a square wave circuit, the ouput of which is converted into square waveforms of pneumatic pressure suitable for driving both assist device and artificial heart.

The pressure levels of the driving waveforms are controlled by means of feedback channels to maintain physiological regulation of the artificial heart's output flow.

A more compact system that could achieve similar regulatory characteristics is also discussed.

INTRODUCTION

Engineering equipment has been used with increasing success to assist the human circulatory system. The human heart is a muscular pump consisting of two main pumping chambers, each with two check valves. The right heart pumps carbon dioxide-laden blood from the body to the lungs at a nominal pressure of 15 mm Hg gage (0.2 N/cm^2 gage). The left heart pumps oxygenated blood from the lungs to the body circulation at a nominal pressure of 100 mm Hg gage (1.3 N/cm^2 gage). The nominal mechanical output power of the heart is about 1.0 watt, but this can increase by a factor of 4 or 6 during exercise.

Electrical and mechanical devices have successfully proven their value to the heart patient. The electronic pacemaker has been developed to excite the heart muscle cyclically when its nervous interconnection becomes ineffective. The heart lung machine can perform the pumping and oxygenation function of the heart and lungs temporarily during an operation.

In the past decade researchers have been attempting to develop small, light-weight pumps for temporarily assisting or permanently replacing the pumping function of the heart. A cooperative program of artificial heart research has developed between NASA-Lewis and the Cleveland Clinic. This work has been sponsored under NASA's Technology Utilization program. An early suggestion made by NASA engineers was to utilize compressed air as an energy transmission medium for the artificial heart. The pneumatic heart designs subsequently developed by the Cleveland Clinic showed considerable promise as compared to their earlier electric motor or electric solenoid-driven pumps. As a result, the mainstream of artificial heart and heart assist development efforts have used the pneumatic pumps.

A number of control system designs have been developed at NASA-Lewis that rely primarily on throttling of pressure from a vacuum and a pressure source to drive the artificial hearts. Some of the earlier control system designs have employed instantaneous servocontrol of the output pressure waveforms. This has resulted in a design of maximum flexibility for doing artificial heart research (ref. 1). Experience gained in the use of these systems has indicated that the quasi-steady-state regulation of the heart over a number of heartbeats is relatively more important than instantaneous waveform control (ref. 2). Elimination of the instantaneous waveform control feature would produce a significant reduction in the complexity and cost of such a control system.

Consequently, the design now receiving attention utilizes appropriately synchronized square waves of output pressure with the levels of pressure and vacuum under feedback control. This system utilizes industrial pneumatic control components. A single-channel version can be used for heart assist applications, while a two-channel version is required for total heart replacement.*

CARDIAC ASSIST PUMP DRIVING SYSTEM

The basic block diagram of the unit as used for an assist pump is shown in Figure 1. The electrocardiogram is fed into an R-wave detector which filters noise with a high degree of reliability and produces an output pulse for each R-wave of the incoming electrocardiogram. The programmer is a pulse circuit which has a square output pulse that is delayed from the R-wave and has an adjustable systolic or ejection duration. The square pulse output of the programmer is converted to pneumatic pressure for systole and vacuum for diastole by means of the pneumatic switching valve. The vacuum and pressure supplies are manually adjusted to obtain optimal filling and ejection pressures at the pump.

---

*Webb, John A., Jr.; and Gebben, Vernon D.: Design and Performance of a Heart Assist or Artificial Heart Control System Using Industrial Pneumatic Components. Proposed NASA Technical Memorandum.

Figure 2 is a block diagram of the R-wave detector circuit designed at Lewis Research Center (ref. 3). The circuit processes the EKG in a sequence of operations which essentially eliminate all components from the input signal except the R-wave. The first block rejects signals that occur equally at the two input leads and attenuates the low- and high-frequency components. The second block rejects low-amplitude signals. Amplitudes that exceed the set level are converted into pulses with durations related to the durations of the original waves. These pulses occur at point D in this figure. The third block rejects short-duration pulses. Since most muscle disturbances produce signals whose waves are shorter than the normal R-wave, this circuit has an exceptional ability to separate the R-wave from the interferences. Results with and without this third block are shown in Figure 3.

These data were recorded during a 2-minute period. The upper trace is the EKG. The first part was recorded while flexing the hand to tense the muscles under the right arm electrode. There was a short period of inactivity before the muscles were tensed again. The middle trace is the result one gets when the pulse-width discriminator is not used. The bottom trace is the output signal from the complete circuit. Interferences that overwhelmed the EKG generally caused the circuit to skip cycles, which is the preferred type of malfunction. When used with special low-noise electrodes, the circuit should be acceptable for clinical applications.

The circuit has been used in heart-assist experiments conducted at St. Vincent Charity Hospital and Cleveland Clinic. Plans are being made to use the circuit in research equipment at the Baylor University and the University of Utah Medical Schools.

Figure 4 is a block diagram of the programmer. The R-wave synchronization signal is used to drive a trigger in the programmer. The output of the trigger can be counted to allow for pulsing on alternate heartbeats or one beat in four. This is thought to have the advantage of preventing chronic dependence of the natural heart on the assist device. The delay block delays systole of the assist pump from the R-wave. This allows

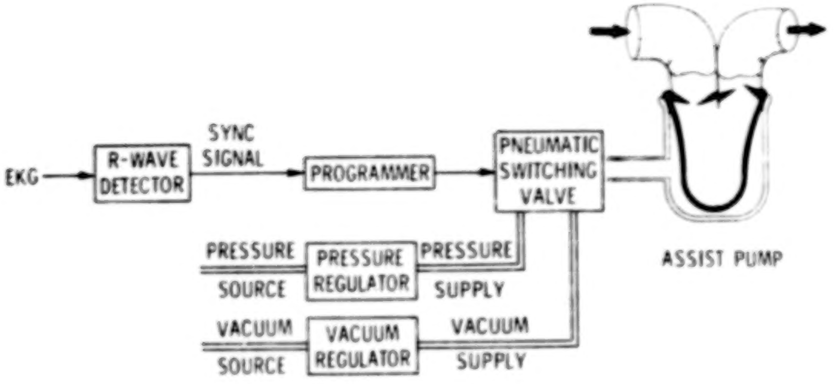

EKG
R-WAVE DETECTOR
SYNC SIGNAL
PROGRAMMER
PNEUMATIC SWITCHING VALVE
PRESSURE SOURCE
PRESSURE REGULATOR
PRESSURE SUPPLY
VACUUM SOURCE
VACUUM REGULATOR
VACUUM SUPPLY
ASSIST PUMP

Figure 1.- Cardiac assist driving system

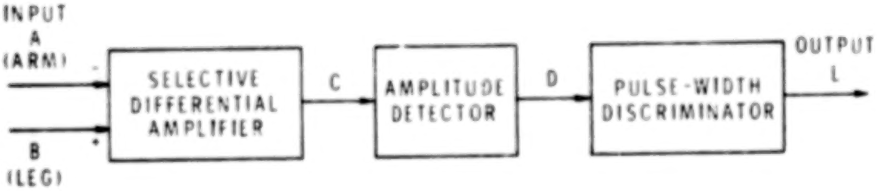

INPUT
A (ARM)
B (LEG)
SELECTIVE DIFFERENTIAL AMPLIFIER
C
AMPLITUDE DETECTOR
D
PULSE-WIDTH DISCRIMINATOR
OUTPUT
L

Figure 2.- Block diagram of cardiac R-wave detector

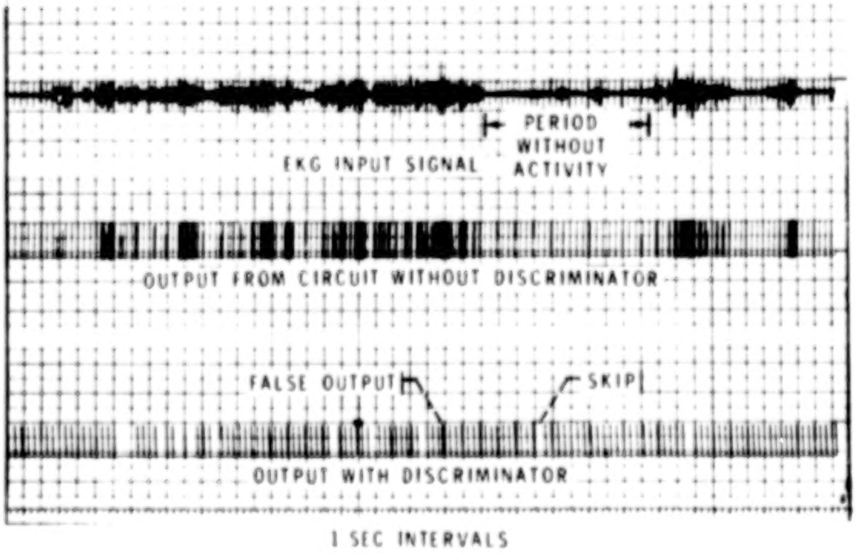


Figure 3.- Comparison of circuits with and without the pulse width disciminator

the assist pump to be operated on natural heart systole or during natural heart diastole which is called counterpulsation. The duration block allows adjustment of systolic duration. The synchronization loss detector will detect a loss of the incoming signal and will switch the unit to a free running generator to maintain pumping. When this occurs an alarm is sounded to inform the operator of the signal loss.

The pneumatic switching valve shown schematically in Figure 5 takes the electrical square wave output of the programmer and converts it to pneumatic pulses. The torque motor alternately caps the control ports of a fluid amplifier, switching its output jet from R_1 to R_2. This switches the spool valve. The output of this valve is then switched between the pressure and vacuum supply. This pulsatile output pressure drives the artificial ventricle.

ARTIFICIAL HEART-DRIVING SYSTEM

This same driving system can be used for a total replacement heart. Of course, the R-wave detector would not be needed, and the programmer would be switched to its own adjustable frequency pulse generator. An additional spool valve driven by the same fluid amplifier used for the pneumatic switching valve will add the second driving pressure. This separate pressure is needed to provide a lower pressure to the right ventricle, while the original spool valve supplies the left ventricle. The systolic pressure and diastolic vacuum for each ventricle now must be adjusted to make the artificial ventricle's output flow a function of atrial pressure. It is believed that this sensitivity is essential to maintain proper cardiac regulation.

Figure 6 gives a plot of ventricular output flow against venous return pressure for the left and right chambers of the natural heart (ref. 4). It is noted that the output flow rate increases as the venous return pressure is increased. This is one of the mechanisms by which the flow rate of blood responds to body needs. It is also responsible for assuring equal long-term blood flow rates for left and right hearts and, therefore, a balance of blood volume between the systemic and pulmonary circulations. In the natural heart simple hydraulics is thought to be responsible for this characteristic. Sensitivities of 5 ℓ/min per 1 mm Hg (1.3×10^{-2}N/cm^2) and 1 ℓ/min per 1 mm Hg (1.3×10^{-2} N/cm^2) must be achieved by the right and left ventricles of the artificial heart, respectively. Due to the high resistance of the artificial valves, this sensitivity cannot be reproduced hydraulically (ref. 5). It is possible, however, to close an electronic feedback loop to affect diastolic vacuum so that the artificial heart's stroke volume can be made to respond properly to changes in upstream pressure (ref.2).

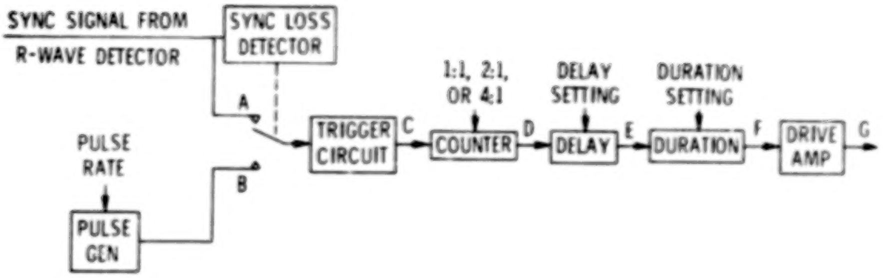

SYNC SIGNAL FROM
R-WAVE DETECTOR
SYNC LOSS DETECTOR
A
B
PULSE RATE
PULSE GEN
TRIGGER CIRCUIT
C
1:1, 2:1, OR 4:1
COUNTER
D
DELAY SETTING
DELAY
E
DURATION SETTING
DURATION
F
DRIVE AMP
G

Figure 4.- Programmer block diagram

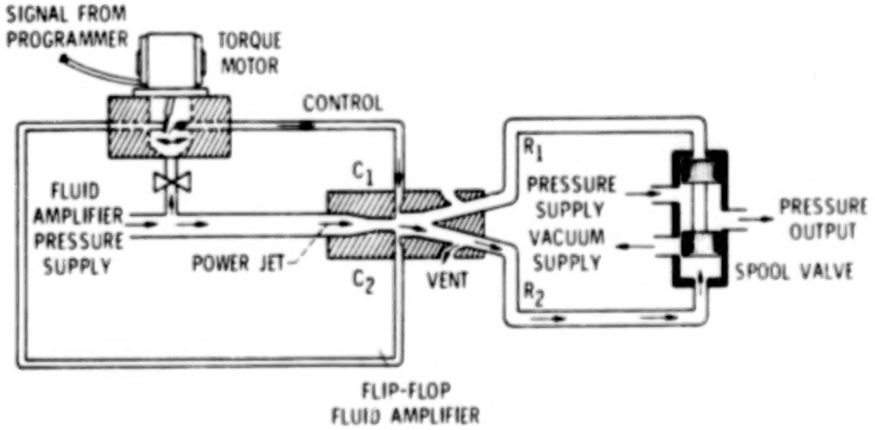

SIGNAL FROM PROGRAMMER
TORQUE MOTOR
CONTROL
FLUID AMPLIFIER PRESSURE SUPPLY
POWER JET
C1
C2
VENT
R1
R2
PRESSURE SUPPLY
VACUUM SUPPLY
PRESSURE OUTPUT
SPOOL VALVE
FLIP-FLOP FLUID AMPLIFIER

Figure 5.- Pneumatic switching valve diagram

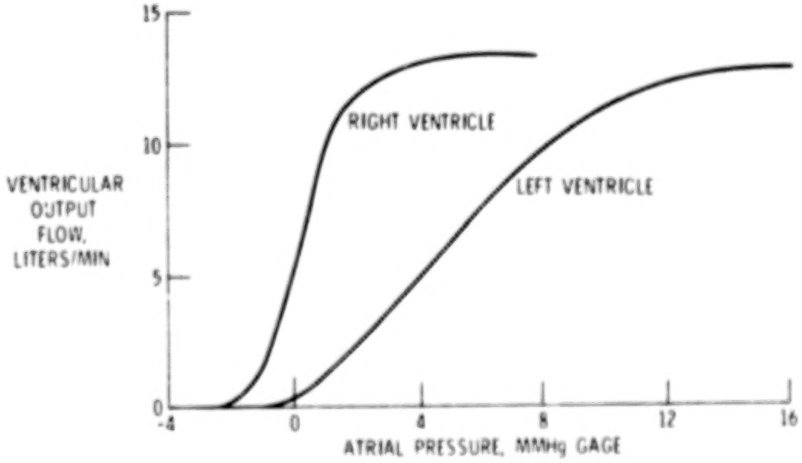


Figure 6.- Normal ventricular output versus atrial pressure for right and left ventricles

By means of a pressure transducer, as shown in Figure 7, the atrial pressure of each ventricle is fed back to control the vacuum used during diastole. This will regulate the end diastolic fill volume. To do this, the pressure signal is averaged with a low pass filter and used as an input signal to an industrial process controller. This controller has an adjustable set point and gain. It drives an electro-pneumatic transducer which has an output pressure proportional to the input current from the controller. Since we are controlling diastolic vacuum, the electro-pneumatic transducer's output pressure must be biased to vacuum. The output of the electro-pneumatic transducer is used as the vacuum supply to the spool valve in the pneumatic switching valve.

If the pressure supply is held constant, the sac will collapse during low flow requirements, since the vacuum has decreased and the pressure is excessive. This collapse will cause increased hemolysis in the pump and must be prevented. To avoid this problem, the pressure must be controlled to maintain the sac volume at some mid-operating position. This is accomplished by using a switch in the ventricle that closes at some mid-position of sac travel. The ventricular volume and switch voltage are plotted in Figure 8 against time to illustrate the way in which the switch indicates excessive or insufficient driving pressure.

By measuring the "on time" of the switch and feeding it back to an industrial controller and electro-pneumatic transducer, the pressure can be adjusted to keep the sac stroking in mid-position as shown in Figure 9.

Figure 10 shows an assembled driving system for one ventricle. The R-wave detector, the programmer, and pneumatic switching valve are the basic bypass components. The industrial controllers are on the panel face with the programmer. Shown next to the pneumatic switching valve is a model of an artificial heart used for testing. The electro-pneumatic transducers and related equipment are mounted in the drawer at the bottom of the console.

In vitro, results were obtained by pumping against a fixed water head of 140 mm Hg gage (1.9 N/cm^2 gage) at a pulse rate of 74 beats/min and systolic duration of 250 msec. By plotting the average output flow of the ventricle against the atrial pressure, the effect of varying vacuum controller gain can be seen in Figure 11. As an indication of atrial pressure sensitivity, the curve from Figure 6 is included for the left ventricle. This shows that normal operation can easily be simulated using this type of control.

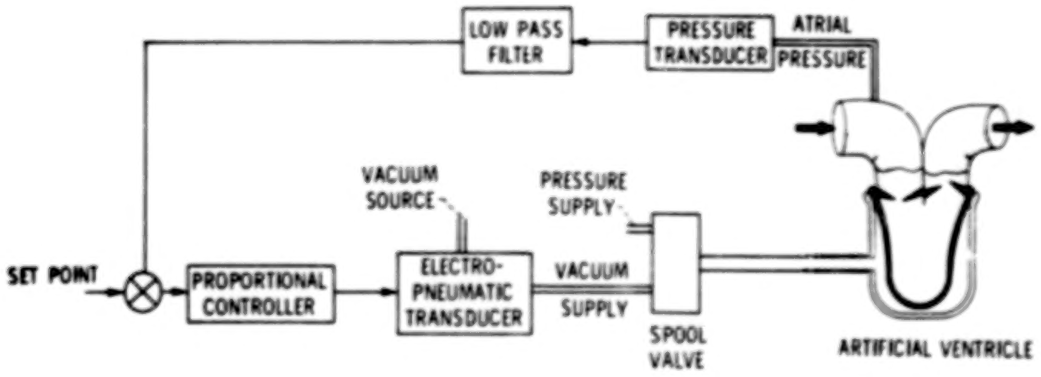

LOW PASS FILTER
PRESSURE TRANSDUCER
ATRIAL PRESSURE
VACUUM SOURCE
PRESSURE SUPPLY
SET POINT
PROPORTIONAL CONTROLLER
ELECTRO-PNEUMATIC TRANSDUCER
VACUUM SUPPLY
SPOOL VALVE
ARTIFICIAL VENTRICLE

Figure 7.- Block diagram of atrial pressure feedback control circuit

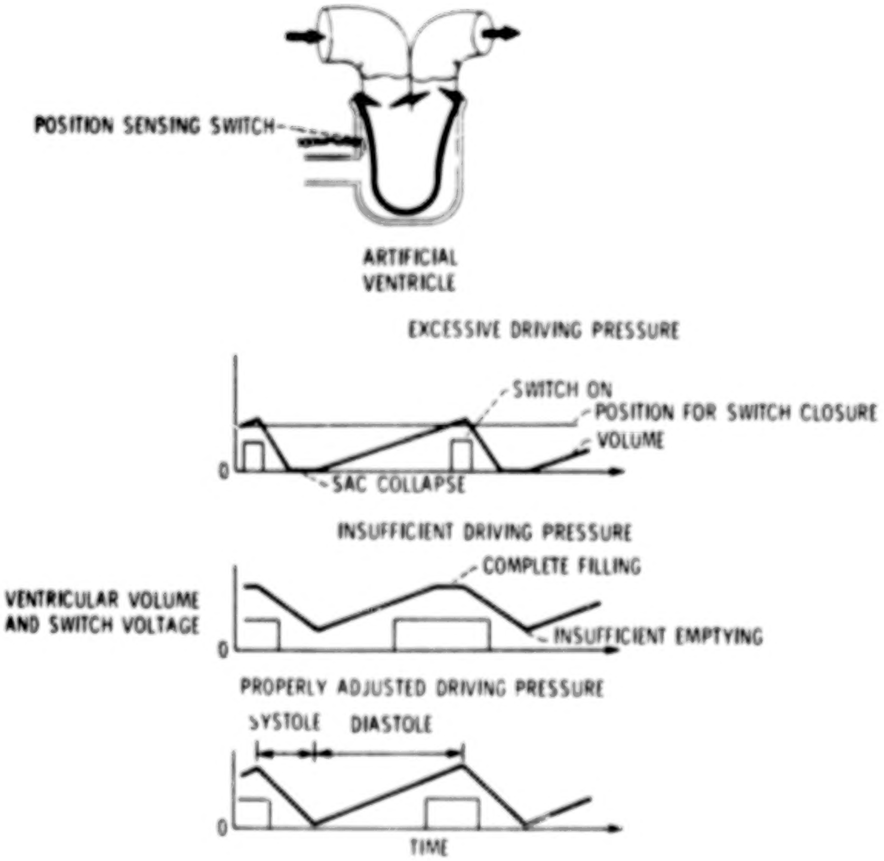

POSITION SENSING SWITCH
ARTIFICIAL VENTRICLE
EXCESSIVE DRIVING PRESSURE
SWITCH ON
POSITION FOR SWITCH CLOSURE
VOLUME
0
SAC COLLAPSE
INSUFFICIENT DRIVING PRESSURE
COMPLETE FILLING
VENTRICULAR VOLUME AND SWITCH VOLTAGE
0
INSUFFICIENT EMPTYING
PROPERLY ADJUSTED DRIVING PRESSURE
SYSTOLE
DIASTOLE
0
TIME

Figure 8.- Ventricular volume detection using switch

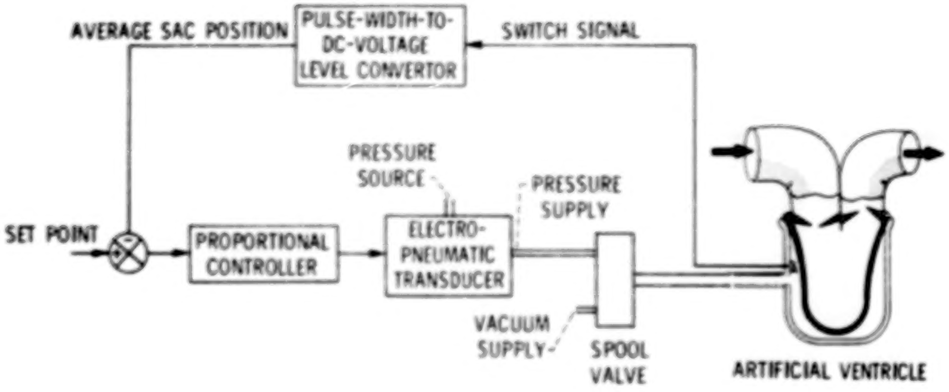

AVERAGE SAC POSITION
PULSE-WIDTH-TO-DC-VOLTAGE LEVEL CONVERTOR
SWITCH SIGNAL
PRESSURE SOURCE
PRESSURE SUPPLY
SET POINT
PROPORTIONAL CONTROLLER
ELECTRO-PNEUMATIC TRANSDUCER
VACUUM SUPPLY
SPOOL VALVE
ARTIFICIAL VENTRICLE

Figure 9.- Block diagram of ventricular position feedback control circuit

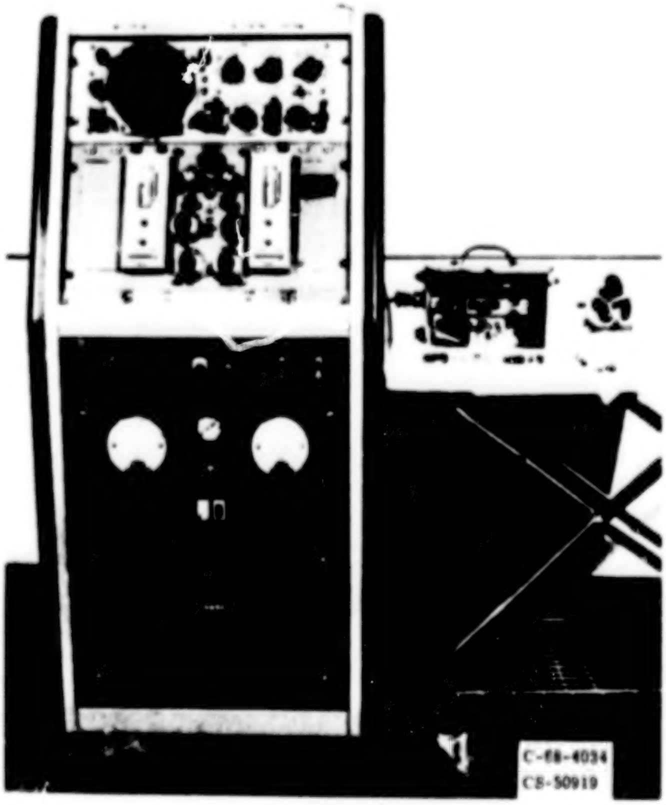

C-68-4034
CS-50919

Figure 10.- Control system console with pneumatic switching valve and artificial heart

FLOW CONTROLLED DRIVING SYSTEM

The artificial heart control system just discussed relies on a pressure transducer for the desired feedback signal. Since the slope of the curves in Figure 6 are extremely steep, variations or drift of 0.5 mm Hg (6.7×10^{-3} N/cm^2) in the transducer can cause a very large change in the cardiac output. One method presently being developed to avoid this problem is to control the heart by measuring the air flow and controlling the cardiac cycle time as shown in Figure 12.

When air flow to the heart becomes zero, the blood displacement into or out of the heart stops. This indicates that the pumping period (systole) or the filling period (diastole) has been completed. The pressure in the air line determines the next operation. For example, at the end of systole when the blood volume in the pump is maximum, the air flow is zero and the pressure is maximum. The sac is then switched to exhaust, which starts the diastolic period. At the end of diastole, the combination of zero flow and minimum pressure in the line provides the proper conditions for switching back to systole.

The cycle rate in this control system depends on the venous and arterial blood pressures. Fast pulse occurs when the venous pressure is high, which fills the pump more quickly; or when the arterial pressure is low, which allows the pump to discharge more quickly.

This control method has the advantage that all the control components are external to the body and the blood pressures. Pressure transducers and limit switches are not installed inside the body. A second advantage is that the blood pressures are limited to the supply and exhaust pressures. Limits result because the air flow stops whenever the blood pressure exceeds the driving pressure or becomes less than exhaust pressure. When the flow stops, the period switches.

The simplest form of this system would drive the heart with a constant supply and a constant exhaust pressure. More elaborate schemes might include pressure amplitudes that vary within the cycle to provide faster flows at the beginning of each period. Also, the amplitude of the driving pressure could be varied as a function of the diastolic time. For example, as the pump rate increased from 50 to 150 pulses/min, the driving pressure could be scheduled to increase from 120 to 180 mm Hg gage (1.6 to 2.4 N/cm^2 gage). The amount of complexity will depend on the basic physiological requirements of the body.

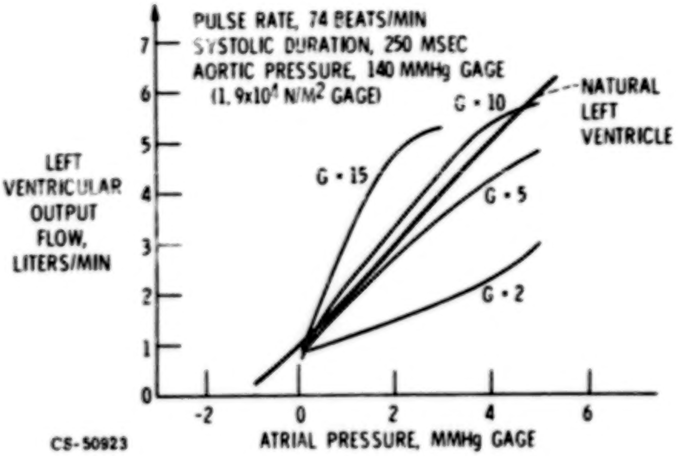


Figure 11.- Left ventricular output flow as a function of atrial pressure

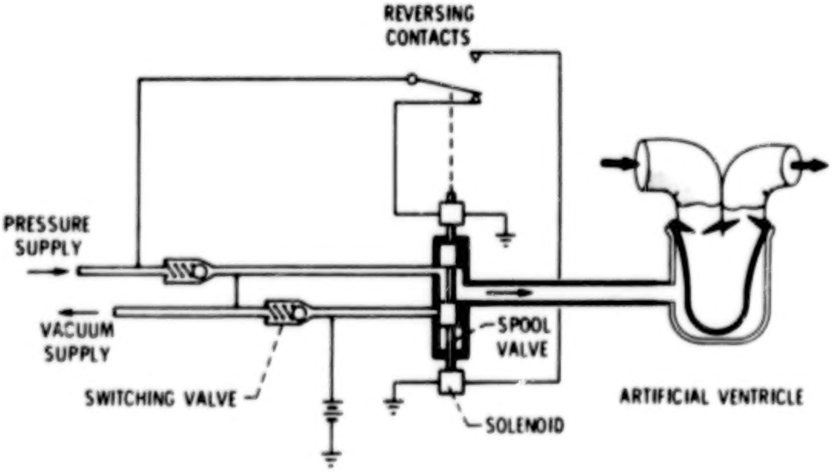

REVERSING CONTACTS
PRESSURE SUPPLY
VACUUM SUPPLY
SWITCHING VALVE
SPOOL VALVE
SOLENOID
ARTIFICIAL VENTRICLE

Figure 12.- Flow controlled driving system

CONCLUDING REMARKS

The assist and total replacement artificial heart driving system provides the medical researcher with a flexible tool for studying various types of feedback control to obtain physiological performance from the artificial heart. The system prevents sac collapse, while maintaining the cardiac output regulation dictated by venous return pressure.

A flow-controlled driving system can be used to obtain a simple system that also can maintain cardiac output by means of venous return. This system has the disadvantage of being inflexible, since the regulatory characteristics of the driving system are dependent on the physical properties of the pump. This may be an advantage from the patient's standpoint, but as a research tool inflexibility limits its scope of usefulness. It does, however, have the advantages of simplicity, small size, and the absence of internal instrumentation.

The design of an artificial heart system must contain control which is adequate to sustain life. It is known that cardiac output flow is a result of complex physiological interactions between central nervous, endocrine, and vascular systems which are not completely understood. All these effects seem to regulate cardiac output through return (inlet) pressure and heart rate. Artificial heart output regulated by inlet pressure (as in this report) has been adequate on short-term animal experiments. Future experiments and long-term survivals may dictate the need for more complex control and dynamics considerations.

REFERENCES

1. Hiller, K. W.; Seidel, W; and Kolff, W. J.: "A Servomechanism to Drive An Artificial Heart Inside the Chest." Trans. Am. Soc. Artif. Intern. Organs, vol. 8, 1962, pp. 125-130.

2. Nosé, Y.; et al: "Respect the Integrity of the Large Veins and Starling's Law." Trans. Am. Soc. Artif. Intern. Organs, vol. 13, 1967, pp. 273-279.

3. Gebben, Vernon D.: "Cardiac R-Wave Detector." NASA TMX-1489, 1968.

4. Guyton, Arthur C.: Circulatory Physiology; Cardiac Output and its Regulation. W. B. Saunders Co., 1963.

5. Kay, E. E.; Zimmerman, H. A.; and Suzuki, A. (assisted by W. S. Griffin and V. D. Gebben): "Prosthetic Valves: Long Term Results." Heart Substitutes. Albert N. Brest, ed., Charles C. Thomas, Springfield, Ill., 1966, ch. 6.

END

DEC 8 1978